Digitized by Arya Samaj Foundation Chennai and eGangotri

JOURNAL OF
THE ROYAL
STATISTICAL
SOCIETY

VOL. 119
1956

CC-0. Swami Bomanji Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

110637
Digitized by Arya Samaj Foundation Chennai and eGangotri

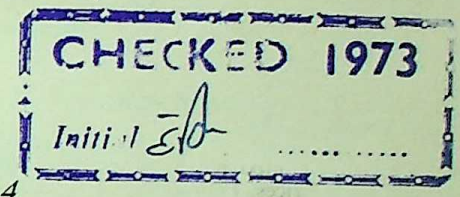
JOURNAL

OF THE

ROYAL

STATISTICAL SOCIETY

SERIES A (GENERAL)



Founded 1834

Incorporated by Royal Charter 1887

Vol. 119—1956

LONDON:
ROYAL STATISTICAL SOCIETY
21, BENTINCK STREET, W.1

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ROYAL STATISTICAL SOCIETY

Patron—HER MAJESTY THE QUEEN

COUNCIL AND OFFICERS—1955-56

Honorary Vice-Presidents

(having filled the Office of President)

THE RT. HON. VISCOUNT SAMUEL, G.C.B., G.B.E., D.C.L.
SIR ARTHUR L. BOWLEY, C.B.E., D.LITT., SC.D., D.SC., F.B.A.
H. LEAK, C.B.E.
THE RT. HON. THE LORD BEVERIDGE, K.C.B., LL.D., D.SC., F.B.A.
ERNEST CHARLES SNOW, C.B.E., D.SC.
THE RT. HON. VISCOUNT WOOLTON, P.C., C.H., D.L., LL.D.
DAVID HERON, D.SC.
THE RT. HON. THE LORD HEYWORTH.
*PROFESSOR A. BRADFORD HILL, C.B.E., F.R.S., D.SC., PH.D.
*SIR RONALD FISHER, F.R.S.
*THE RT. HON. THE LORD PIERCY, C.B.E.

* Also Member of Council.

President

PROFESSOR E. S. PEARSON, C.B.E.

Vice-Presidents

PROFESSOR D. G. CHAMPERNOWNE.
D. J. FINNEY, SC.D.
PROFESSOR M. G. KENDALL, SC.D.
SIR GEORGE MADDEX, K.B.E., F.I.A.

Other Members of Council

PROFESSOR R. G. D. ALLEN.
P. ARMITAGE, PH.D.
PROFESSOR G. A. BARNARD.
PROFESSOR M. S. BARTLETT, D.SC.
B. BENJAMIN, B.SC., F.I.A.
H. CAMPION, C.B., C.B.E.
STELLA V. CUNLIFFE.
H. E. DANIELS, PH.D. (Apptd. October 1955).
IRIS DOUGLAS.
E. C. FIELLER.
R. F. FOWLER, C.B.E.
R. F. GEORGE.
R. G. GLENDAY, M.C.
DAVID HERON, D.SC.
JOAN E. KEEN.
W. P. D. LOGAN, M.D.
PHILIP LYLE (died July 1955).
C. A. MOSER.
E. D. VAN REST.
W. F. SEARLE.
G. W. SEARS.
SIR ROBERT SHONE.
J. STAFFORD.
J. R. N. STONE, C.B.E.
L. H. C. TIPPETT.
L. T. WILKINS.
F. YATES, SC.D., F.R.S.

Executive Committee

THE PRESIDENT.
THE EX-PRESIDENT.
THE VICE-PRESIDENTS.
THE HON. SECRETARIES.
THE HON. TREASURER.
B. BENJAMIN.
DAVID HERON, D.SC.
PROFESSOR A. BRADFORD HILL, C.B.E., F.R.S.

Honorary Treasurer

R. F. FOWLER, C.B.E.

519.905
R81J
V.119

Honorary Secretaries and Editors

R. F. GEORGE — IRIS DOUGLAS
B. BENJAMIN (appt. October 1955)

पुस्तकालय विज्ञान भवन
विषय सं०..... 519.905
पुस्तक सं०..... R827 V.119
आ. पं. सं०..... 4257
गुरुकुल कांगड़ी

Honorary Foreign Secretary

R. F. GEORGE

Secretary

UNA M. CROKER

Librarian

J. A. KING

Bankers

ROYAL BANK OF SCOTLAND (DRUMMOND'S BRANCH), CHARING CROSS, LONDON, S.W.1

Offices

21, BENTINCK STREET, LONDON, W.1
(Telephone: WELBECK 7638)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CONTENTS

VOL. 119—YEAR 1956

Part 1, 1956

	PAGES
The Statistician and the Planning of Field Experiments. By D. J. FINNEY	1–17
Discussion on the Paper	17–27
Causal Inference from Observational Data, A Review of Ends and Means. By HERMAN WOLD	28–50
Discussion on the Paper	51–61
Some Aspects of the Structure of Indian Industry. By VERA ANSTEY and RUSSI J. TARAPOREVALA	62–82
The principles of Factor Analysis. By M. G. KENDALL and D. N. LAWLEY	83–84
Regulations of the Medical Section	85
Reviews of Books	86–98
Statistical and Current Notes	99–101
Obituary:	
H. W. Methorst	102
Charles Stewart Orwin	103–104
Statistical and Economic Articles in Recent Periodicals	105–111
Additions to the Library	112–121
Periodical Return: Bank of England Weekly Return, 1955	122–123

Part 2, 1956

	PAGES
Some Aspects of the Geometry of Statistics: the Use of Visual Presentation in Understanding the Theory and Application of Mathematical Statistics. The Inaugural Address of the President, Professor E. S. PEARSON, C.B.E., delivered January 18th, 1956	125–146
Proceedings of the Meeting	146–149
The Analysis of Business Concentration: A Statistical Approach. By P. E. HART and S. J. PRAIS	150–181
Discussion on the Paper	181–191
Richard Jones: A Reappraisal. By R. GLENDAY	192–193
Wholesale Prices in 1955. By THE EDITOR of *The Statist*	194–198
Banking Statistics. By W. MANNING DACEY	199–209
Reviews of Books	210–220
Statistical Notes	221–224
Statistical and Economic Articles in Recent Periodicals	225–230
Additions to the Library	231–237

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Part 3, 1956

	PAGES
United Kingdom Indices of Wholesale Prices, 1949–1955. By H. S. PHILLIPS	239–276
Discussion on the Paper	276–283
The Television Audience in the United Kingdom. By B. P. EMMETT	284–306
Discussion on the Paper	306–311
Some Theory of Index Numbers. By YOU POH SENG	312–332
A Geometrical Derivation of the Analyses of Covariance and Variance. By C. P. COX	333–335
Note on the History of Sampling Methods in Russia. By S. S. ZARKOVIC	336–338
Reviews of Books	339–351
Statistical Notes, Additions to Library, Periodical Returns	352–370

Part 4, 1956

	PAGES
Estimating Expenditure, Saving and Income from Household Budgets. By DOROTHY COLE and J. E. G. UTTING	371–387
Discussion on the Paper	387–392
The Cost of Road Accidents. By D. J. REYNOLDS	393–408
The New Zealand Agricultural Sample Survey. By W. B. TAYLOR and D. V. P. CLEMENT	409–424
Some Theory of Index Numbers—II. Maintenance of Continuity by the Assumption of Proportionality of the New Index Number to the Old Index By YOU POH SENG	425–455
Note on a Visit to Moscow for the Third Soviet Mathematical Congress. By M. S. BARTLETT	456–458
Annual Report of the Council	459–476
Proceedings of the One Hundred and Twenty-Second Annual General Meeting	477
Reviews of Books	478–487
Statistical and Current Notes	488–491
Obituary:	
John Wishart	492–493
Arthur Avery Adams	493–494
Mrs. J. R. N. Stone	494
Statistical and Economic Articles in Recent Periodicals	495–499
Additions to Library	500–505
Periodical Returns: Revenue of the United Kingdom, 1955	506
Index to Vol. 119 (1956)	506–508

NOTICE

THE Council of the Royal Statistical Society wish it to be understood that the Society is not responsible for the statements or opinions expressed in the Papers read before the Society or published in its *Journal*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Vol. 119.] [Part I

Journal of the Royal Statistical Society

SERIES A (GENERAL)

PART I, 1956.

The Statistician and the Planning of Field Experiments

By D. J. Finney

University of Aberdeen ; Agricultural Research Council Unit of Statistics

[Read before the Royal Statistical Society, November 16th, 1955, Professor M. G. Kendall, Vice-President, in the Chair]

Contents

	PAGE
1. Introduction	1
2. The nature of experimental agriculture	2
3. Statistics and scientific research	3
4. General policy	3
5. Particular aspects for experimental study	4
6. Pilot experiments	4
7. Design of experiments	5
7.1 Selection of treatments	5
7.2 Specification of plots	6
7.3 Specification of measurements	7
7.4 Allocation of plots to treatments	8
7.5 Replication	12
8. Amount of experimentation	12
9. Conduct of experiments	13
10. Recording	13
11. Statistical analysis	14
12. Reporting and interpreting	14
13. Summaries of groups of experiments	15
14. Concluding remarks	16

1. *Introduction*

Whether or not they are directly concerned in designing experiments, most statisticians are now aware of the ideas underlying experimental design. Certainly any statistician engaged on problems of scientific and industrial research must be familiar with the principles, and must have a detailed knowledge of those classes of designs most frequently needed in his own field. Nevertheless, much research still fails to benefit fully from the statistical outlook on experimentation.

The statistician can give much more help to an experimental programme than is involved in selecting a Latin square, randomizing, and performing an analysis of variance on the results. I believe that the time is ripe for surveying the whole impact of modern statistical thought on experimentation, not with any intent to demarcate an enlarged statistical empire, but rather to show the regions in which statisticians must be able and willing to collaborate with their colleagues if the best use is to be made of all effort put into experiments. This paper is an attempt at one part of such a survey.

In agriculture, as in medicine, much important research can be conducted in the laboratory. In medicine, the final appeal must be to the effect of a new treatment on human subjects; so in agriculture, no new crop variety, fertilizer, method of cultivation, or other treatment can be

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

established until its field performance has been studied and compared with alternatives. I shall confine my discussion to these field experiments. I have elsewhere (1955, Chapter IX) briefly considered similar ideas in a broader context, but only by focusing attention on one particular branch of science can I introduce the amount of detail that seems desirable. Moreover, in addition to my personal interest in agricultural research, I have the excuse that the great development of experimental design during the last thirty years began primarily to meet the needs of agriculture; even today the statistical aspects of experimentation are probably more advanced and more widely understood here than in most other sciences. With some translation of terms and some shift of emphasis, however, the ideas that I shall put forward are relevant to many other branches of research. I take as my title *The Planning of Field Experiments* rather than *The Design of Field Experiments*, because I want to examine the responsibilities of statisticians in relation to the whole structure and conduct of research programmes, from the first realization that an experimental study of a problem may be necessary to the completion of reporting. Experimental design is an important part of statistical activity in any such programme, but it must not be the whole interest of the statistician concerned: one aim of this paper is to encourage statisticians to be less timid in assessing their own responsibilities. On so broad a front, I must express many personal opinions and I shall therefore freely use the first person singular.

2. *The Nature of Experimental Agriculture*

Two principles must be recognized as applying to practically all research in agriculture:

(*a*) Experimentation is limited by economic factors;
(*b*) Experimentation is sequential.

In an applied science, the economic limitation is frequently encountered directly: the funds available for a project are determined by financial administrators (who, in different instances, show different degrees of attention to the estimated importance of the outcome of the research). The economy of a situation is not always so clearly expressed in monetary values, however, and may instead be measured by a limited availability of man-power, of time, or of land on which experiments are to be conducted, though all of these are roughly convertible into money values for the sake of comparison with one another and with any estimated value of the outcome of the research. Even in the purest of research, analogous limitations apply. For example, the total amount of some new material for study may be severely restricted, or a special piece of apparatus may at present be the only one of its kind. Moreover, a research worker untroubled by any such restrictions must still appreciate that each day of his life spent on a particular investigation means one day less for other competing interests. Whatever the precise nature of the limitations and however great may be the difficulty in expressing them quantitatively, there is clearly an obligation to seek some optimal utilization of resources; in general terms, any experimental programme should be planned with a view to maximizing the value of the results for the specified expenditure of money, materials, or time.

Inevitably a sequential element enters all agricultural research, even though the programme is not explicitly expressed according to the special techniques of sequential analysis. The search for improvement in agricultural practice through the discovery and adoption of better varieties of seed, better types and amounts of fertilizer, better cultivations, and so on, always continues. No absolute best is ever attained; to some extent the aim is always changing, because of new discoveries in other branches of research, alterations in markets and in prices, development of new machinery, and other trends that modify the definition of "best". Even under otherwise rigid conditions, the unpredictability of weather would always prevent any absolute optimal being specified. Nevertheless, the programme adopted at any time should rest upon all knowledge gained in the past on the same or related topics. Rules cannot be formulated wth the exactness of an orthodox sequentially designed experiment, but the design and conduct of each new experiment should depend to an important degree upon the accumulated evidence of the past. Imperfections of reporting by other workers can make examination of this evidence a great labour, and a balanced judgment is needed in deciding how long the start of new work can be delayed by it. One consequence of a more unified approach to the whole subject of field experimentation ought to be improved reporting and greater accessibility of information.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

3. *Statistics and Scientific Research*

All research that involves a quantitative assessment of results necessitates some form of statistical treatment, unless the quantities measured are determined, by assignable and measurable causes, with such accuracy that any residual inexactness is of no importance to the investigator. The analysis may be no more than the calculation of means or a simple graphical presentation of measurements; alternatively, it may involve long and complex computation. Whatever the methods employed, no conclusions are firmly based without some attempt to distinguish between the effects of factors that are under experimental study and the uncontrollable variations usually grouped under the name of "experimental error". This is certainly true of research by agricultural field experiments: plots of land alike in treatment and in every respect that the experimenter can control may give very different crop yields, so that a comparison between different treatments always needs to be considered in relation to an assessment of its experimental error.

For ease of discussion, I distinguish between *the statistician* and *the experimenter*, using the latter term for any research scientist whose work requires that he shall conduct experiments or cause them to be conducted.* Many experimental scientists to-day have sufficient knowledge of statistical principles and techniques for most of their own problems, though even they are likely occasionally to seek assistance on special points from the professional statistician. Others, either from inclination or because they lack this knowledge, ask much more frequently for statistical help from the specialist; yet others, perhaps, could benefit as much from this help but do not realize their need of it.

That a high proportion of those whom I have broadly classed as experimenters should have a knowledge and appreciation of statistical principles is of great value to the progress of their science, but perhaps not so much because they are thereby enabled to do their own statistical work as because they understand better what needs to be done and what conclusions can reasonably be drawn. The extent to which any one man should be his own statistician, even when he has the requisite knowledge, depends largely upon circumstances and personal inclination; what is important is that both he and his statistical colleagues should have sufficient understanding of one another's disciplines to avoid the errors that can result from imperfect communication of ideas. The whole training of a statistician is directed to making him think quantitatively, and his experience should bring him special facility in handling quantitative information. Too often he is regarded, and even regards himself, as concerned solely with the technical processes of calculation and perhaps with the narrower aspects of choice of experimental design. In fact, his training and experience should fit him for *participation in every stage of research in which quantitative thinking is relevant*. His part is neither that of a servant of other research workers, nor that of their master. He should be a collaborator, at some stages playing a major role, at others being one of many, and at others perhaps little more than a spectator seeking to improve his general knowledge of the problems under investigation in order to increase the value of his own contributions.

Yates (1953) has vigorously advocated a broader interpretation of the responsibility of statisticians in agriculture. In amplifying and illustrating my own ideas by reference to field experimentation, I shall have much in common with his general survey; I shall, however, be more detailed, and shall consider the statistician's duty in relation to each of ten phases of a research programme.

4. *General Policy*

Agricultural research is an applied science, though it can range from *ad hoc* investigations into problems of immediate importance to long term investment in work that will add to the general body of knowledge. Whatever may be true of associated pure sciences, such as botany, chemistry and entomology, the justification for spending time, money, and other resources on any aspect of agricultural research must be a reasonable hope that it will eventually lead to some improvement in agriculture, whether by higher yields, by healthier crops, by better working conditions, or by some other change.

Hence at some stage the economy of research needs to be considered in relation to the national economy. In order to determine what investment of effort in any proposed programme ought

* Of course, no reference is intended to the Scientific Officer and Experimental Officer classification of the Scientific Civil Service.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to be made by those who ultimately control the resources of agricultural research, the probable return from the programme must be assessed. That the economic aspects of general research policy are not always given prominence does not alter the facts. Any decision to increase research into weed control, or milk production, or even agricultural statistics, presumably springs from a belief that, in the long run, the gain will more than balance the total cost to the community. To express these ideas quantitatively is always difficult, especially as any advance assessment of the consequent benefits must be a guess rather than a trustworthy estimate. Nevertheless, this problem is in part statistical, and a statistician should contribute to the analysis of the situation with the aid of experience from analogous research in the past.

Moreover, economic evaluation of a research programme must often use extensive knowledge of current farming practice. Conventional and long-established censuses, especially annual returns from agricultural holdings, will supply some of this. In addition, if the merits of changes in specific aspects of farming practice are under consideration, information will be needed on what that practice is. How much fertilizer are farmers using, and in what way? What is the extent of weed growth in various crops? Such questions can be answered satisfactorily only by carefully planned sample surveys, an excellent example being the series of *Surveys of Fertilizer Practice* conducted from Rothamsted (Yates, Boyd and Mathison, 1944; Church, 1952). The need for the statistician in planning and analysing both the censuses and the sampling is obvious. Perhaps less generally accepted is his part in synthesizing estimates of the economic consequences of suggested changes in terms of costs, increased production, and net gains. Rough guesses of the type of result that may emerge from a research programme may be combined with survey information to indicate what benefits might accrue from the research, so assisting the choice between different lines of enquiry. Later, when the technical and administrative problems of implementing recommendations based upon research have to be tackled, more reliable estimates of the effects of suggested changes per unit area will be available from experiments; if the potential benefits to the whole agricultural economy, or to particular sectors thereof, are to be assessed, information from surveys must again be used.

5. *Particular Aspects for Experimental Study*

General policy will determine that research must be undertaken on certain broad subjects—methods of weed eradication, the adaptation and development of a crop new to a region, or the improved utilization of fertilizers. Particular aspects of these problems must be chosen for detailed study. Is it more expedient to seek to control weeds by cultivation or by chemical sprays? Which features of the growing of a new crop, such as breeding new varieties, selection for high yield or disease resistance, modification in manuring or in cultivation, are likely to be most susceptible of improvement? If the benefits from fertilizers are to be increased, is the most hopeful line of enquiry heavier use on certain crops, better techniques of application, or the production of new chemical compounds with special properties?

Obviously these questions must be largely answered by the technical experts, in the light of their knowledge, the abilities and interests of the research workers who are available, and the availability of materials, machinery, etc. In part, however, the answer depends upon quantitative assessment of the present state of knowledge from previous work or from similar investigations in other countries or on related crops. Here again the statistician has a part to play in assembling and evaluating the evidence. Insistence on this point may seem almost an anxiety to create a job for the statistician; however, I have known instances of series of experiments on the spacing of a crop or on the time of application of fertilizers being planned quite without regard to published information on the same subject from a neighbouring region of the country. Whether or not the experimenter acts as his own statistician on this may depend upon the size of the task, but some attempt to digest the old work should be regarded as an essential preliminary to the beginning of the new (see § 13).

6. *Pilot Experiments*

Occasionally, the lack of reliable evidence on questions relating to a series of experiments may justify expenditure of time and resources on preliminary *pilot experiments*. These may help to bridge the gap between laboratory and field (e.g. by pot experiments) or may relate directly to field

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

conditions but on a small scale. They can be useful in showing the practicability of certain treatments, in training staff for unusual types of investigation, in indicating the order of magnitude of treatment effects, and even in giving quantitative information on the merits of alternative sizes of plots and on components of variance. On the other hand, only a fairly large experimental programme can stand the diversion of any appreciable amount of effort into special studies of experimental technique. Not only will the statistician wish to use results from pilot experiments in his subsequent work on designs, but also he should be prepared to advise on the extent to which the experimenter is wise to delay the start of his main researches by using part of his resources in pilot studies.

7. *Design of Experiments*

Questions relating to the design of an experiment can usefully be considered as of five types:

7.1. Choice of treatments for comparison;
7.2. Specification of the units to which treatments are to be applied;
7.3. Specification of the measurements to be made on each unit;
7.4. Rules by which units are allocated to the various treatments;
7.5. Number of replications of any one factor or of the whole experiment.

7.1. *Selection of Treatments*

Some experimenters, and even some statisticians, appear to think that the latter should simply design experiments to suit the treatments proposed by the experimenter without venturing upon criticism. In reality, an important duty for the statistician is a critical examination of the logical structure of the treatments in an experiment. His appreciation of the quantitative economics of experimentation should inform his views on whether a proposed experiment is capable of answering the most important questions that interest the experimenter.

A superficial study of agricultural experiments over the last fifty years might suggest that the trend to fewer and simpler treatments, apparent about 1920–30, had since been reversed, the sets of treatments now used being sometimes as complex as in the more ambitious experiments at the beginning of the century or earlier. The first effect of R. A. Fisher's influence on design was to make experiments consist of many replicates of a few simple treatments. Since then, advances in agricultural knowledge and the fuller development of Fisher's ideas have enabled results from seemingly very complicated combinations of treatments to be interpreted. Fifty years ago, treatments for an experiment were often chosen for their separate individual interests to the experimenter, seldom on any carefully integrated plan. Today we know how to do things better, although unfortunately experiments with almost haphazard choice of treatments can still be found. Any large experiment demands a carefully elaborated pattern of interrelationships of treatments; the design determines the form of statistical analysis that is appropriate, and, in particular, the structure of the treatments determines what questions can usefully be asked of the results. The statistician's problems are closely related to the subsequent allocation of plots to treatments (§ 7.4). I shall mention here, however, two examples of more general issues on which he must be prepared to express his opinions with knowledge and firmness (Finney, 1953).

I have known an experimenter to claim that he merely wished to compare two treatments, such as two methods of weed eradication or two levels of a fertilizer. He has argued that he can omit untreated controls, because he already has evidence that his treatments are qualitatively beneficial and he merely seeks to compare the alternative forms. In this simple situation, doubtless most statisticians would insist that, unless the experimenter knows his other treatments to be *in all circumstances* beneficial, untreated control plots be included. The variability of the conditions under which crops are grown is such that, on a particular occasion, neither method of weed eradication might increase crop yield: without a control yield to indicate this, the results of the experiment might be interpreted as evidence that the two methods were equally good.

A less trite example is that of the experimenter who wishes to estimate the quantity of fertilizer per acre required for maximum yield of a crop. To achieve this with reasonable precision would be difficult, because the regression curve relating yield to amount of fertilizer is almost always very flat over a wide range of heavy dressings. However, the experimenter should be brought to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

see that, for most practical purposes, the estimate he needs is *not* the dressing of fertilizer giving maximum yield but the most economic dressing. Practical advice will aim at telling the farmer what amount of fertilizer can be expected to maximize the difference between its cost and the value of the crop; experiments should be planned to estimate this, and to study the response curve over a range of amounts so that recommendations can be modified to suit changed economic conditions.* Some experimenters will argue that they must try to estimate conditions leading to a maximum yield, because in times of national emergency and scarcity every possible effort would have to be put into maximizing agricultural production. This misses the point. From the point of view of constructing an approximate statistical model, an emergency is a violent change in the economic structure in which all prices and values move far from their normal levels. A scheme of experimentation planned for estimating the optimal fertilizer application at current prices can scarcely be expected to be very good for estimating the optimal under an entirely different set of values. Moreover, in so far as an emergency is likely to involve labour and fertilizer scarcities as well as to place an increased value of the crop, it may easily have the effect of making the optimal lower instead of nearer to that giving maximum yield. The experimenter must decide whether his object is to obtain information relevant to present conditions, whether he is estimating for some hypothetical future, or whether (and to what extent) he will compromise the one aim in order not to neglect the other entirely.

7.2. *Specification of Plots*

In agricultural field experiments, the experimental units capable of differentiation in respect of treatment are usually "plots" comprising small areas of land and the crop plants growing thereon. The dimensions of plots must be largely influenced by convenience and by the practicability of the agricultural operations to be performed upon them. Nevertheless, within the range of what is admissible, evidence on components of variance, from past experiments and uniformity trials on the same crop under comparable conditions, assists the statistician to judge how plots of different dimensions would compare in respect of variability (§ 6). He may be called upon to advise on the relative merits of a few fairly large plots or a greater—but not proportionately greater—number of small plots; his answer must depend upon some assessment of the replication that these will permit and the likely precision of treatment comparisons in relation to the relative costliness and difficulty in use of the alternative plot sizes.

To speak and write glibly about this somewhat unexciting aspect of experimental design is easy, but when we need to choose a plot size for a particular experiment the procedure may be less simple. Not only must dimensions of the harvested area be considered, but the need for guard areas of unharvested crop and possibly for pathways between plots complicates the discussion. Also, one may be sure that the ideal specifications will vary from place to place and from season to season. Fortunately, the optimal is likely to be fairly "flat", so that the loss consequent upon deviation from the best possible dimensions will often be small. For bush and tree crops the problems are generally more difficult than for annual plants; moreover, the number of past experiments that can be used for guidance is smaller, and any dimensions chosen in a particular instance have to be satisfactory for a much longer time. If I dismiss the subject briefly here, that is because of the difficulty of discussing it further in general terms and not because I underestimate its importance. Twenty years ago, a considerable amount of effort in agricultural research was apparently directed at comparisons of plot size and related matters, sometimes perhaps to the exclusion of questions of true agricultural importance! Without advocating any overemphasis on the study of experimental technique, I do suggest that the great increase in field experimentation in recent years may make economically desirable rather more study of this subject than it seems to

* There can be no inflexible rule, but the following is often a reasonable guide; it assumes that, over the range of levels that is important, the response curve can be represented, very approximately, by a quadratic equation, and the rather more realistic Mitscherlich equation, $Y = A\,(1-10^{-kx})$, is usually sufficiently close to this. First guess the optimal level. Then conduct the experiment with zero and 2 or $2\frac{1}{2}$ times the guessed optimal as extremes, together with one or two additional levels equally spaced between them. To use much higher amounts is less efficient, because the experiment then studies sections of the response curve that discriminate poorly between levels of application; to introduce many different levels instead of three or four usually involves putting excessive effort into studying the shape of the response curve while determining the optimal less precisely.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

receive: possibly statisticians could make more organized attempts to collate and classify the evidence from large numbers of experiments. There can never be rigid rules, but summaries of past experience, such as that of Garner and Weil (1939), can be useful for the guidance of future research.

7.3. *Specification of Measurements*

Naturally the choice of measurements to be made on every plot depends primarily upon the interests of the experimenter. He may be concerned chiefly with studying the yield of the crop, or the degree of insect infestation, or the density of weed growth, and the main variate that must eventually be submitted to statistical analysis will be a measurement appropriate to this. Even here, however, statistical considerations enter, especially when the measurement must be based upon a sample of a plot. For total crop yield this will not often happen, but study of qualities such as percentage protein or sugar in the crop, or of pest and disease infestation, will almost always have to be based upon sampling within the plot. In planning a suitable sampling scheme, the participation of the statistician is obviously essential.

The statistician can also help the experimenter to choose suitable units of measurement. When yield is the primary interest, there is relatively little difficulty, though sometimes the alternatives of yield per unit area and yield per plant, or of yields in various quality or size classes, have to be considered. In research on pests and diseases, there is more scope for argument. Is the appropriate measure of the effect of treatment the number of insects per unit area, or per plant, or per leaf? Must any account be taken of different sizes or stages of development of the insects, or of their depth in the soil? Is some index of the extent of visible damage, such as number of leaf punctures or area of leaf affected, more suitable (or more easily obtained and no less suitable)? Is ultimate crop yield still an important criterion for the success of treatments, since this is what directly concerns the farmer? These are only to a small degree statistical questions, but the decisions taken may greatly affect the nature of the conclusions that can be drawn from the experiment. From his expert knowledge of the analysis that will follow, the statistician should be able to put clearly before the experimenter some of the good and bad points about the various possibilities, so aiding him to make a choice consistent with obtaining useful information yet not requiring excessive or unnecessary labour.

Often the experimenter himself will wish to make several measurements on each plot, either because all are equally important or because some give subsidiary information bearing upon his chief interest, such as measures of secondary products or of quality of the crop. Field experimentation is expensive, but additional measurements can be obtained relatively cheaply on an experiment already sanctioned for one purpose. The well-informed statistician will have ideas on whether such measurements will repay their cost, either directly in relation to the main study or because the experiment is necessarily providing materials relevant to other research.

The choice of measurements depends so closely on the aims of the experiment that frequently a need for clarification of terminology appears. Words such as "larger", "best", "fertility" are not always as simple as a dictionary might suggest, and, if the statistician is to give the appropriate help at this stage of planning (and at others), he may need to insist that the experimenter defines his meaning more exactly.

Another category of additional measurements (on the crop or on the soil) consists of those that have little direct importance but that may prove valuable as concomitants in a covariance analysis. The only justification for measuring them may be that they can increase the precision of main comparisons more than would employment of an equivalent effort in increasing the replication of the experiments. For example, an experimenter on sugar-beet can relatively easily count or estimate the number of plants per plot. He might be interested in plant density itself, or might have his own reasons for wishing to discuss yields adjusted to equality of plant density; if not, plant counts are not worth making unless a covariance analysis increases the precision of comparisons between more important measurements to a greater extent than would a similar expenditure on increasing replication or improving the experiment in another way. This relates to measurements planned when the experiment is begun but the decision to measure a concomitant such as plant number is sometimes taken at the time of harvest, because unforeseen circumstances have caused it to vary excessively and thus exceptional reasons make adjustment for it likely to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

increase precision. The statistician has a responsibility for assembling evidence and utilizing it to advise experimenters on the expected gains from concomitants.

7.4. *Allocation of Plots to Treatments*

The two principles of *randomization* and *replication* remain at the heart of all good experimental design, though both are now—quite correctly—interpreted in a manner that sometimes mystifies the non-statistician. To an audience of statisticians, I hope I need say little about the first. No satisfactory substitute for randomization in ensuring freedom from bias and validating the estimation of error seems possible; most attempts to analyse unrandomized experiments consist in applying various adjustments and then asserting that what remains can be analysed as though there had been randomization. The need for replication, at least of individual factors though not necessarily of every combination of factors, will scarcely be denied by statisticians, though some experimenters still wish to lay down only one plot of each of several treatments. Usually this is defended by the argument that only "demonstration" or "observation" plots are wanted: closer inquiry may disclose that what is to be demonstrated is far from being well-established as a desirable practice, or that what is to be observed is unlikely to be interpretable without information on variation between plots of the same treatment! Moreover, if any large number of plots is required, subdivision into blocks is practically essential in order to keep the error variance for important comparisons of treatments reasonably small; the price paid is that differences between blocks correspond with unimportant comparisons.

Why should the number of treatments in any one experiment, and therefore the number of plots, be large? Fisher, Yates, and others have stressed the need for factorial experiments, in agriculture as in other branches of science, but the matter is sufficiently important to bear re-examination. Farming inevitably involves many factors. These fall into three main groups. First there are factors whose states or levels are directly controlled by the farmer, such as types and quantities of fertilizer, varieties of seed, and dates of field operations. Secondly are factors over which the farmer has no control: climatic factors and the inherent physical and chemical properties of the soil come into this group. Thirdly are factors that might be controlled if there were any evidence that they affected the crop, but that are ordinarily dismissed as irrelevant: the day of the week on which the land is ploughed, the phase of the moon at sowing time, and the colour of the tractor-driver's shirt. The groups are not entirely independent; for example, weather will influence the farmer's choice of dates for various operations, and his fertilizer applications are intended to compensate for deficiencies inherent in his soil.

Whenever a farmer grows a crop, he does so under particular conditions in respect of all factors. He chooses the conditions in respect of some factors, and needs to select not merely the best for each independently but the best combination: the best choice of fertilizer will depend upon the variety of seed, and this in turn will depend upon the time at which he wishes to sow and to harvest his crop. Moreover, in using evidence from experiments, he needs assurance either that they were conducted under conditions similar to his own with reference to uncontrolled factors, or else that the effects of the controlled factors are not affected by the uncontrolled. He chooses a variety known to be suited to his climate, applies fertilizers deemed optimal for that variety on his own soil, and does not worry whether his tractor-driver has a shirt of the same colour as one who worked on the experiments that guide him. In short, he needs information on the magnitudes of increases and decreases in crop to be expected from changes in some of the many factors that constitute his practice of husbandry, and the extent to which these magnitudes depend upon other factors. So far as the controlled factors are concerned, it is precisely this information that is contained in—to use the jargon of experimental design—the set of main effects and interactions in a factorial experiment, even though the summary of these that comes most naturally to the statistician may not always be best suited to practical needs.

Suppose that an experiment compares three different inter-row spacings for sugar beet, using only one level of nitrogenous fertilizer. In applying the results, *either* inferences must be restricted to crops receiving exactly the same amount of nitrogen, *or* the effect of changes in spacing must be assumed independent of the amount of nitrogen, *or* the effect of changes in spacing for other levels of nitrogen must be predictable with sufficient accuracy from the estimates given by the experiment. Agricultural science is still largely empirical. The theory that would justify the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

third type of inference rarely exists, although such predictions may sometimes be aided by evidence from other multifactorial experiments. The second possibility, though no doubt acceptable if "colour of shirt" were to replace "amount of nitrogen", is a type of argument known to be false in so many situations that to adopt it in one less familiar is very unwise. The first possibility is safe but of small value, since the limitations would need to apply in respect of so many factors simultaneously. Only by inclusion of different amounts of fertilizer in the one experiment can the difficulty be removed. Then, if yield differences associated with the spacings are found to be dependent on nitrogen level, the effects at certain levels are directly estimated and those at intermediate levels can be predicted by interpolation; if not, there is direct evidence in support of an assumption of independence even at levels other than those tested.

The need for including several factors in one experiment is plain, but how far can it be pursued in practice? I am convinced that many experiments are still planned with inadequate appreciation of what can be done by factorial design, confounding, and fractional replication. Often opportunities are lost of adding factors to a projected experiment at practically no cost yet with great gains in the information obtained and the breadth of its applicability. An experiment on one factor at two levels would be of poor precision unless it had about 8 replicates, say in randomized blocks of 2 plots. Two such experiments on different factors might be combined, using all combinations in 8 blocks of 4 plots: the loss from the larger blocks would be offset by having 16 replicates instead of 8 so far as the main effects are concerned, and in addition the interaction of the factors is examined. In this experiment, 21 out of 31 degrees of freedom are used in estimating error variance and only 3 are assigned to treatment comparisons, By introducing a third factor and arranging the design as four randomized blocks of 8, there will be some loss in precision from the larger blocks but now 7 d.f. are measuring treatment comparisons; 21 d.f. are still estimating error, so that only one-quarter of the effort of the experiment is spent directly on the treatments. Important though the error is, this is extravagant, for an error mean square based on 21 d.f. is not intrinsically much more valuable than one based on say 10 d.f. Either one or two further factors can be introduced, without any change in block size, by confounding and using interactions as a measure of error (see below). A suitable design estimates 5 main effects and 10 two-factor interactions, using only 13 d.f. for error, so that now slightly more than half the 28 d.f. within blocks relate to treatments. Not only can the design be saturated with factors in this way, but fractional replication provides a method of supersaturation. The 2^5 design can be made into a 2^6 (6 factors instead of 5), provided that one interaction of two factors can be sacrificed; the design then uses 20 d.f. for main effects and two-factor interactions, only 8 d.f. for error.

If a 2^n factorial experiment with r-fold replication on each factor has an error mean square s^2 with f degrees of freedom, the precision of any one effect estimated from it is

$$\frac{r(f+1)}{2s^2(f+3)}$$

(Fisher, 1951). Hence, if the experiment permits a main effects and interactions to be estimated free of any confounding, the quantity

$$\frac{a\,r(f+1)}{2\,s^2(f+3)}$$

may be regarded as an index of the information extracted from it. For the main effects and two-factor interactions from 2^1, 2^2, 2^3, 2^4, 2^5, 2^6 designs, this takes the values

$$3.2/s^2,\ 22/s^2,\ 44/s^2,\ 72/s^2,\ 105/s^2,\ 131/s^2;$$

even allowing that the first requires only half the space of the others and that smaller blocks are likely to lead to smaller values of s^2 for the first two, the advantages of the 2^5 and 2^6 designs appear substantial. Of course, this index must not be employed uncritically, for it makes no distinction in value between main effects and interactions; nor has any account been taken of the occasional need for estimation of three-factor or even higher order interactions.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

One difficulty with confounding and fractional replication is that the possibility of measuring certain interactions, judged in advance to be unimportant, must be sacrificed. Moreover, in an experiment that is saturated with factors, variances must be estimated from other interactions. This is less serious than might at first appear, since interactions of high order (i.e. involving four or even three factors simultaneously) are likely to be relatively small, so that the comparisons that should measure them are in fact measuring little more than error variance. The statistician will choose a design that, taking into account existing knowledge of interactions, minimizes any ill-effects of confounding and of the absence of error estimates based on true replication. Of course he cannot be sure that he and the experimenter will guess the unimportant interactions correctly. However, the experimenter must decide between, say, testing factors A, B and C with no confounding and with true replication but at fixed levels of D and E, or testing all factors simultaneously and risking other difficulties of interpretation. To lose some interactions by confounding is scarcely less satisfactory than to restrict the basis of inference by omitting factors from the experiment. To inflate the estimate of error a little, by including in it interactions that are not truly negligible, is often not very serious when at the same time inferences about the first three factors are being extended so as to relate to several states of D and E. In a final comment on a paper to this Society twenty years ago (Yates, 1935)—a paper whose lessons for experimental design are still incompletely accepted—Yates said: "In general all arguments which attempt to show that factorial design is ineffective or misleading when interactions exist contain their own refutation, since it is only possible to assess interactions by experiments of the factorial type." At the same time, an experiment must not be so loaded with factors that too few degrees of freedom are left for satisfactory estimation of error, for which about 6 d.f. seems a reasonable minimum.

The inclusion of many factors is limited by the size of experiment that can be carried by the institute and research workers concerned. Confounding and fractional replication become easier to arrange, and their disadvantages become less, as the size of the experiment increases, but the number of factors can seldom be as large as one would like in a thorough investigation of a group of problems. Here again the declining importance of higher order interactions is a protection.

The subject is complicated, and nothing I have said must be regarded as an uncompromising advocacy of multifactorial single or fractional replicates for every purpose. My concern here is to insist that the statistician must know the potentialities of factorial design and be able to put clearly before an experimenter the possible gains from adding factors. If the design chosen is multifactorial, the statistician will help in finding the best scheme of confounding; if not, he should ensure that the experiment is conducted on few factors for positive reasons rather than for lack of thought. Inclusion of many factors, especially if this involves adopting a design that does not replicate every combination of levels, undoubtedly makes the interpretation of an experiment more difficult; although this difficulty often has its compensations in respect of the greater information obtained, there are situations in which the risk is too great and the experiment should be kept down to one or two factors. For example, in experiments for which the same plots are to be used during many years, whether with a rotation of annual crops or with a perennial crop, to predict that certain interactions will remain unimportant throughout may be unwise. Many complications unsuspected at the beginning can arise in long-term experiments, and to have two or more plots of each treatment combination is a considerable safeguard; sometimes this is conveniently achieved by starting different replicates in successive years. Despite the many merits of *successful* multifactorial fractional replicates, a simpler design that gives trustworthy results on a limited front is better than a scheme too ambitious for the circumstances of the experiment. Though I have emphasized the importance of considering the possibility of incorporating additional factors into an experiment, I must recognize that the statistician not infrequently has to resist pressure from an experimenter who wishes to include many factors in unsuitable circumstances.

A serious limitation in research on annual crops is that one experiment can refer only to one location in one year. Must its inferences be limited to that place and that assembly of meteorological conditions? If so, its value is indeed small! Something can be gained by repeating the experiment at the same place in several years, either on neighbouring sites or by building it into a long-term rotation experiment on the same plots. However, if factors whose levels the experimenter cannot control are to be at all adequately covered, there must also be replication in space. Experiments of the same or similar design can sometimes be laid down at several research farms as part of a co-operative scheme.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The results of agricultural experiments are so susceptible to changes in soil and climate that something more is needed. These few experiments may indicate the factors and interactions likely to be of chief practical importance, but they are less satisfactory for quantitative estimation of the gain that would follow widespread adoption of particular recommendations; this flaw will be increased by any tendency for soils and standards of husbandry on research farms to differ from the average of the regions they serve. A much more widely dispersed series of experiments is needed, preferably on normal farm land under normal management. For obvious reasons, such experiments must be very simple and must use only a few plots. For many years, several countries of Northern Europe have recognized this need: in Denmark, for example, extensive programmes of simple experiments on farms have been organized and reported since the early years of the century. Unfortunately, rather little thought has been given to planning such series in a manner that would serve the needs of research as well as acting as "demonstrations", and too often the experiments have been considered individually instead of as elements of an integrated programme. Interesting progress has been made recently in India, especially in Bihar State where the energy of Dr. H. N. Mukerjee has enabled several thousand experiments on the main crops to be completed in each of the last six years. He attempts to select sites at random, from all land in the State or a subdivision thereof that is to grow a particular crop.* The possibility of using several sets of treatments in one series of experiments is now being explored. For example, in a large series of fertilizer experiments on Indian food crops subsidized by the U.S. Technical Co-operation Administration, the following scheme is one of several that have been used. Each experiment has six plots only, and the treatments (in a new random order for each experiment) are one of the three sets:

(i) None, P, N_1P, N_2P, N'_1P, N'_2P;

(ii) None, P, N_1P, N_2P, N''_1P, N''_2P;

(iii) None, P, N'_1P, N'_2P, N''_1P, N''_2P;

P denotes unit dressing of phosphate, N, N′, N″ refer to three different sources of nitrogen, and subscripts denote single and double dressings. The three sets, used in equal numbers, can be combined to give estimates of the effects of phosphate, the response curve for nitrogen, and the relative merits of the nitrogenous fertilizers. The experiments are a type of fractional replication; no systematic classification of designs or of methods of statistical analysis and interpretation has yet been prepared, though Panse and Sukhatme (1953) have given a fairly brief discussion.

The statistician needs to be able to employ both the multifactorial designs and the series of individually simple but collectively complex small experiments. Both types may be needed in the same programme, the one for fundamental research and the other to provide a broad basis for practical recommendations. If experiments are to continue on the same plots for several years, he must remember that experimenters' ideas change with the accumulation of evidence from other research. He will therefore be wise to try for a flexible design that may permit treatments to be changed and new ones to be introduced, without serious damage to the orthogonality of the whole. For example, Latin squares can be modified to Graeco-Latin, and unsaturated factorials can have new factors put in. For this and other reasons, especially if any confounding is involved, he will look most favourably on designs in which all factors have the same number of levels, especially 2^n and 3^n. Detailed discussion of the technicalities, however, must be left to the specialized books on design.

I cannot discuss the important class of balanced and partially balanced incomplete block designs used for comparing many varieties of a crop, or for other situations in which there is no true factorial pattern in a fairly large number of treatments. These are a type of confounding, but new problems arise in their design and analysis. Questions relating to the optimal number of treatments or varieties to include have as yet scarcely been formulated, but may become a concern for the statistician.

* To secure a strictly random selection is impracticable, because of the inaccessibility of some localities and the unwillingness of some (very few) cultivators to co-operate. However, the area actually sampled will correspond fairly closely to that most likely to be influenced by any advisory measures subsequently based upon the experiments.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Despite my insistence (§ 2) that agricultural research is sequential, I have said nothing about individual sequential designs: I doubt whether these can be used to advantage in field experiments: successive experimental periods must be a year apart, errors are large relative to the effects to be estimated, and climatic and other uncontrollable factors alter substantially from year to year the circumstances under which experiment proceeds. It seems as though the sequential development of knowledge and of ideas for further research must arise from non-sequential unit experiments. However, new thoughts on sequential design may change this outlook, and the agricultural research statistician should not overlook the potentialities of sequential techniques, for which uses have already been suggested in other branches of agriculture (e.g. Fisher, 1952).

7.5. *Replication*

For any projected experiment, the experimenter should have some idea either of the magnitudes of differences between treatments and the precision with which he wishes to estimate these or of the minimum differences that he is interested in detecting as statistically significant. From his knowledge of the experimental errors usually encountered (cf. § 7.2), the statistician can assess what replication is desirable in order to achieve satisfactory precision. Cochran and Cox (1950, Chapter 2) have given an excellent elementary discussion of this problem, but there is undoubtedly room for further extension of realistic theory.

Decisions on replication cannot be divorced from other aspects of design. Not only are replications and plot size conflicting interests in respect of the total area occupied, but the replication judged necessary for one factor determines or affects the number of additional factors whose inclusion should be considered.

8. *Amount of Experimentation*

The experimental designs discussed in the last section fell into two categories, the large and complex that are practicable only at research stations or other centres at which skilled staff can be supplied, and the very simple intended for wide dispersion in large numbers on the land of ordinary farmers. The number of experiments of any one complex design that can be carried out in a single year is severely limited. Moreover, the more complex the design the more likely is it that the accumulation of information and the changing views of experimenters will make some modification desirable in the next year.

For the simpler experiments, a question of immediate relevance is: How many of any one design should be conducted? The greater the number of experiments, the more precisely will the results averaged for a region indicate the right practice for general adoption in that region. The rate of gain from increasing the number, however, decreases as the total number increases, and eventually falls below the rate of increase in the cost of the experimental programme. Progress has recently been made with the problem of determining the optimal number of experiments, under one or two different specifications of the questions to be studied.

For example, Yates (1952) examined the problem of assessing how many experiments can most profitably be employed in estimating the optimal level of application of a fertilizer. Clearly the number should depend upon the total area to which the fertilizer is eventually to be applied, as well as upon the value of the crop and the cost of a single experiment. Yates took into account the average reduction in net profit to be expected when the optimal is estimated with a known variance instead of being itself known exactly, and determined the number of experiments that would minimize the sum of that quantity and the cost of the experiments. Quite large deviations from this number do not greatly affect the net gains, so that rough advance estimates of all the quantities involved can be a useful guide. Prices and costs vary from year to year, but at least some objective basis for deciding the number of experiments is here indicated. In practice, a rather larger number can usually be justified. One reason is that the estimated optimal level of fertilizer will presumably continue to be used for several years. Secondly, a slightly larger investment in the experimental programme may help to produce evidence on which different fertilizer recommendations for different parts of the whole region can be based.

Grundy, Rees and Healy (1954) have recently considered experimentation directed to deciding between two discrete alternatives: Shall an established practice be continued or shall a new proposal

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

be recommended in its place? They discuss how evidence from a few experiments can be used to determine the economically desirable additional number that should precede the final decision. This problem is technically more difficult, but the same general principles are applied.

The problems mentioned in this section are extreme simplifications; for example, no distinction has been drawn between experiments in one season and experiments spread over several seasons. This phase of the statistician's duties is likely to be influenced by research on decision functions and related topics during the next few years. Enough has been said to show that the statistician has a distinctive contribution to make to the decision on how many experiments of a particular type shall be conducted, even though he is not yet armed with a complete theory to represent mathematically the practical questions that must be faced.

9. *Conduct of Experiments*

The only stage of the research programme at which the statistician's technical skills have no special role is the actual conduct of the experiments. I suggest, however, that he will not waste his time, bias his judgment, or harm himself in any other way if he occasionally soils his hands by assisting in routine field work: the experimenter will often be glad to have intelligent, though unskilled, help in sampling, weighing, and measuring. Moreover, he should familiarize himself with the appearance of different crops and different types of experiments. In this way, he will improve his knowledge of the ends to which his research is directed and the means of attaining these. He will come to appreciate some of the technical difficulties encountered by the experimenter, and so to devise experimental designs best suited to the peculiarities of particular crops and treatments. He will learn to speak in the language of the experimenter. He may even come to be regarded by his non-statistical colleagues as a man who knows something about agricultural research, and not merely as a mathematician parasitic on agriculture.

For these and related reasons, I believe that statisticians closely concerned with the details of experimentation should be located at research institutes, where they are in close contact with the day-to-day progress of the work in which they are collaborators, even though others whose prime concern is with general policy may need to be more closely tied to administrative centres. The young recruit to agricultural statistics whose first years of experience are spent in a central government office is unlikely to develop a well-informed and sympathetic habit of co-operation with the experimenter. If his work is to consist of the design, analysis, and interpretation of experiments, he should be placed in an environment that helps him to understand the aims of those whom he helps and the practical difficulties that beset them; no appeal to the administrative convenience of having him situated centrally and equally accessible to all should be regarded as necessarily an over-riding factor. When he has become reasonably experienced in the rapid appreciation of experimental problems, however, a man based on one university or research institute may be able to give statistical advice to other centres.

10. *Recording*

The systematic collection of records of experiments and their arrangement in tables of standard pattern are a safeguard against errors and omissions and an assistance to all statistical analysis. Though he must not expect to have everything his own way, the statistician should give a lead in planning good recording. He must remember the importance of having a recording system that can be used simply in the field under possibly adverse conditions. At the same time, he will wish to minimize the transcription of original records into "fair copies" for analysis, and to provide as many opportunities as possible for cross-checking on any preliminary calculations and conversions of units done before the records reach him. I have come to be very suspicious when neatly typed results of experiments are submitted to me for analysis; too often they contain gross mistakes, not least because experimenters and typists, unaware of their own fallibility in copying large tables, omit to impose thorough checks.

Not only must records be systematically collected, but also they must be full and detailed enough for all purposes connected with the interpretation of the experiment. Both the experimenter and the statistician should consider, in the light of their past experience, what information about the circumstances and conduct of the experiment, as well as what measurements on the crop (§ 7.3), are worth recording and what can be omitted in the interest of a reasonable economy of effort.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

11. *Statistical Analysis*

Here there can be no doubt about the duty of the statistician. In § 10, I have urged him to ensure that stringent checks are made on the accuracy of records that reach him; I should add that for him to permit arithmetical or transcription errors during the course of an analysis under his control is inexcusable. An experimenter who knows something about statistical techniques may choose to analyse his own experimental data. If he does, he should set himself the same high standards of integrity, accuracy, and critical investigation that he would rightly demand of a professional statistician.

The principles of statistical analysis of field experiments are well standardized. Excellent and well-known standard text books, and a wide range of individual papers, make this perhaps the easiest of the statistician's tasks, even though for complex experiments it may be a laborious process. I can scarcely write more on this topic without expanding this one section into a book. Let me emphasize that, although statistical analysis may be the one aspect of the research programme for which the statistician is solely responsible, it must never become his only interest and activity.

12. *Reporting and Interpreting*

The statistical analysis of field experiments frequently consists largely in analysis of variance, calculation of standard errors, performance of tests of significance, and the preparation of formal statements about estimates of treatment effects. A tabulation of these operations is not in itself a report. The analysis of variance, for example, is merely a routine procedure in the reduction of the results, normally having no more place in a final report than a discussion of the size of plot or of the make of tractor employed. Tests of significance and estimates of the effects are the scaffolding around which a report must be constructed, but their meaning needs expression in language intelligible to the consumers of the report. Personally, I deplore in any branch of research the presentation of a report on an experimental programme followed by an independently prepared statistical appendix. Certainly in agriculture no distinction should be drawn between the *practical* and the *statistical* interpretation of an experiment. The statistician must play his part in translating his findings into agricultural terms, and in providing an integrated series of conclusions.

My views on the reporting of the analysis of variance and allied computations may seem extreme. Let me emphasize that I refer only to reports concerned with the translation of research findings into agricultural policy and practice. Even here, there may be exceptions, just as there will be experiments in which information on the tractor is essential to interpretation. His work on a particular piece of agricultural research may stimulate the statistician to develop new statistical theory or technique, and in publishing this he may wish to show many details of analyses of variance or other computations. Such publications may lead to improvements in experimentation, but are entirely different in character: practical conclusions from experiments are likely to be less effectively implemented if their reporting is combined with discussion of experimental technique or statistical theory.

The statistician experienced in agricultural research can contribute to the drafting of reports much more than the technical points that are his special concern; he is less likely than the experimenter to have preconceived ideas on experimental results, and can usefully employ his objective outlook on the whole research. Moreover, his mathematical training should have accustomed him to an especially clear habit of thought in the logical development of an argument. Lest I should seem to make extravagant claims, let me state that I do not for one moment suggest that report writing be turned over entirely to statisticians! The ideal division of responsibilities in any instance must depend upon the individuals involved: I am anxious only to combat the notion that the statistician must refrain from comment on any part of a report not explicitly relating to statistical analysis. A man schooled in the precise language of mathematics should see the need for some mastery of the written word in any scientific report and, without becoming pedantic, should train himself to avoid infelicities of style, ambiguity, obscurity, unnecessary repetition, and bad grammatical construction. Above all he should aim at simplicity, the use of concrete expressions that avoid all vagueness, and the choice of words exactly suited to particular concepts. On these matters there is always scope for argument and no statistician need set himself on a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

pedestal, but, by giving conscious thought to the development of his own style, he will increase the value of his collaboration and increase the chance that others will read and understand what he writes. Indeed, I doubt whether a statistician really understands his own analyses unless he can put himself in the place of the reader for whom his report is written, and can describe their general outline and meaning intelligibly.

13. *Summaries of Groups of Experiments*

In agricultural research today, advance rarely comes from a single dramatic experiment. More commonly, improvements in agricultural practice can be based only on the critical appreciation of evidence from a large number of experiments which are not necessarily all of the same type or even all from the same organized programme. From time to time a synthesis of information on some particular topic is needed, for which purpose all evidence must be extracted from the relevant past experiments, collated and critically examined, and used as a basis for general recommendations. The thorough examination of past experience that ought to be the foundation of any further research on the same topic also requires the critical summarizing of experiments.

A striking example of this process was the examination of evidence on the fertilizer requirements of the major British food crops, made by Crowther and Yates (1941). Despite many years of experimentation on a great variety of sites widely dispersed over the country, no clear pattern of relationship between experimental evidence, advisory services to farmers, and current farming practice had emerged. The experiments had been organized, conducted, and reported by many different authorities, no one being responsible for co-ordinating the results. The grave food shortage of the time, and the difficulties and dangers of ensuring imports of food or of fertilizers, made urgent the planning of a policy that would bring optimal fertilizer dressings to the most important crops. The authors sifted the very heterogeneous information, discarded aspects of experiments that were of no direct importance to fertilizer needs (but never exercised a subjective judgment on the inclusion or exclusion of evidence on fertilizer responses), and finally, by a careful averaging of responses over regions and soils, produced a clear and consistent picture of fertilizer responses. This became the foundation of policy on fertilizer imports and use.

Though less vitally important to the national welfare, general economic considerations clearly demand that new experimental programmes also should rest on critical analysis of all existing relevant information. Whether previous experiments were bad, or good but inadequate to give clear answers to all the questions, they should not be ignored. Too often an individual research worker, a research institute, or even a body charged with broader responsibility for agriculture research, initiates an investigation apparently without even being aware that others have studied the same problem in the same or similar crops in the past.

A summary must be not merely a hotch-potch of results from many different sources with comments on their concords and their discords, nor even a précis of conclusions previously presented by those who reported the constituent experiments. It must be a closely integrated interpretation of the whole body of information, discussed critically and leading to recommendations as detailed as the evidence can justify. New statistical analyses will be needed, and these may be more difficult than analyses of single experiments because their methods are not standardized. The task requires considerable statistical experience and skill—in saying which, I am not forgetting the agricultural knowledge that is essential. In the simpler instances, a statistician well-informed on agriculture may himself be capable of surveying past experiments and summarizing their indications on some special topic; more often, close collaboration with other agricultural scientists will be the best guarantee of a valuable report.

Participation in such projects will teach a statistician a great deal about the designing and reporting of experiments. He will learn a proper respect for complete honesty of reporting, and a refusal to ignore or pass over without exact statement experiments that have not given statistically significant differences between treatments. I have known the value of a series of experiments to a more extensive summary to be completely destroyed by the discovery that, in addition to some experiments reported in full numerical detail, there were others of the same series merely described as having failed to give significant results; any averaging based upon the experiments reported fully is then liable to serious bias.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

This preparation of summaries of groups of experiments is an end and a beginning. It is the final phase of an experimental programme, putting results into a form that can be used by farmers and those who advise them, yet it is also the starting point for planning new research. It should answer some of the questions raised under §§ 4 and 5, and it will be a prelude to a new application of the procedures of §§ 4–12.

14. *Concluding Remarks*

When once I lectured less formally on this theme, one of my hearers suggested that, if the statistician did so much, nothing was left for the experimenter or anyone else! Though I have confined my attention to the statistician, nothing is further from my intention than to present him as the most important figure: neither statistician nor statistical analysis is a panacea for all the ills of experimentation. Every statistician associated with a programme of scientific research needs to remind himself frequently that not only is his work valueless unless built upon good experimental practice, careful operation of equipment, painstaking observation, and exact recording, but that these fundamentals—and above all the inspiration of whoever first thinks of a new line of research—are far more important to scientific progress than his own contribution. Good research can be done when an initial inspiration is examined by common sense and a critical appreciation of observations, even though refined statistical techniques are excluded. I am convinced that a statistician can often make good research still better. If the inspiration and scientific ability of the experimenter are lacking or are misapplied, however, no amount of statistical skill can remedy the mediocre research that must ensue.

The statistician is an adviser who, from expert knowledge, offers facts, figures, and opinions that will influence policy, but who does not himself have ultimate responsibility for determining and executing that policy. He collaborates closely with experimenters in designing and analysing experiments, but in general he will not himself be responsible for initiating the research. He suggests designs that, in the light of statistical principles, seem optimal, but their adoption must always be limited by their practicability within the resources of the experimenter and his staff. He analyses, and the relevance of his interpretation depends upon the skill of the experimenter in the execution of his work. He may summarize experiments from many sources and produce from them recommendations on agricultural practice: if he is wise, he will do so only in collaboration with those who have better claim to be in touch with agricultural needs, unless his research experience has given him some expert knowledge of agriculture as well as of statistics. In short he will not be permitted, nor will he himself aspire, to change his proper confidence in the value of his own contribution into an assurance that agricultural research and statistical skill are synonymous.

Even with this insistence on the modesty of my statistician, I may seem to have drawn him too large. Few statisticians could ever be expected to deal adequately with all my ten phases of a research programme. Frequently, some phases are beyond the competence of any statistician, either because the requisite theory does not yet exist or because my demands that advice must rest upon information from past experiments cannot be met when that information does not exist in any suitable form. Nevertheless, we can usefully look at the whole range of statistical activity in the planning of field experiments and endeavour to see the interrelations of the parts, for so we become aware of gaps in knowledge and flaws in practice. Doubtless my statistician is a composite figure, whose functions must be performed by several men of more specialized interest and abilities; that division of responsibility however, is relatively unimportant and must be adjusted in accordance with the special interests and abilities of the individuals involved.

I do not regard this paper as in any way final: my aim is to provoke thought, discussion, and even disagreement, and I earnestly hope that my hearers will translate what I have said into terms relevant to their own fields of statistics. I have outlined an ideal pattern of statistical service to research, knowing well that it cannot at present be completely realized. Nevertheless, the preparation of this paper has brought home to one statistician his many inadequacies. I believe that we can profitably argue about the ideal, adding new responsibilities for the statistician that I have omitted, modifying (but I suspect not subtracting from) those that I have included. Perhaps all of us who are directly concerned in agricultural research can attempt to approximate to our own versions of the ideal. To others, I suggest that, by thorough examination of the part that should

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

be played by statistics in the planning of field experiments, we may eventually learn much about the proper function of statistics and statisticians in other activities of our community.

Acknowledgment

For sixteen years I have enjoyed a great amount of guidance on agricultural experimentation from Dr. F. Yates. Though I cannot now be sure exactly which ideas in this paper I owe directly to him, undoubtedly a large part derives from what he has taught me. In acknowledging my debt, however, I must make clear that Dr. Yates is not committed to agreement with any of my opinions.

References

CHURCH, B. M. (1952), "Recent trends in fertilizer practice in England and Wales", *Emp. J. Exp. Agric.*, **20**, 250–270.
COCHRAN, W. G. & COX, G. M. (1950), *Experimental Designs*. New York: John Wiley.
CROWTHER, E. M. & YATES, F. (1941), "Fertilizer policy in war-time", *Emp. J. Exp. Agric.*, **9**, 77–97.
FINNEY, D. J. (1953), "Response curves and the planning of experiments", *Ind. J. Agric. Sci.*, **23**, 167–186.
—— (1955), *Experimental Design and its Statistical Basis*. Chicago University Press.
FISHER, R. A. (1951), *The Design of Experiments*, 6th ed. London: Oliver & Boyd.
—— (1952), "Sequential experimentation", *Biometrics*, **8**, 183–187.
GARNER, H. V. & WEIL, J. W. (1939), "The standard errors of field plots at Rothamsted and outside centres," *Emp. J. Exp. Agric.* **7**, 369–379.
GRUNDY, P. M., REES, D. H. & HEALY, M. J. R. (1954), "Decision between two alternatives—how many experiments?", *Biometrics*, **10**, 317–323.
PANSE, V. G. & SUKHATME, P. V. (1953), "Experiments in cultivators' fields", *J. Ind. Soc. Agric. Stat.*, **5**, 144–160.
YATES, F. (1935), "Complex experiments", *J. R. Statist. Soc. Suppl.*, **2**, 181–247.
—— (1952), " Principles governing the amount of experimentation in developmental work", *Nature*, **170**, 138–140.
—— (1953), "The wider aspects of statistics", *J. Ind. Soc. Agric. Stat.*, **5**, 109–118.
—— BOYD, D. A. & MATHISON, I. (1944), "The manuring of farm crops: some results of a survey of fertilizer practice in England", *Emp. J. Exp. Agric.*, **12**, 164–176.

DISCUSSION ON DR. FINNEY'S PAPER

Dr. YATES: It gives me very great pleasure to propose the vote of thanks to Dr. Finney for his most interesting paper.

I particularly like the approach he has adopted in the paper. He has stepped back from the details of experimental design on which, as we all know, he is an expert, to look at the broader aspects and decide what part statisticians have to play in these broader aspects. All who have had contact with experimental work know that it is these broader aspects which present the most difficult problems. It is relatively easy to become expert in the art of designing individual experiments and in the art of analysing experiments when they are completed. It is much more difficult to decide what kind of experiments should be undertaken, what questions should be included, how the treatments should be chosen so as to throw most light on the selected questions, and so on. This is a field to which the statistician has much to contribute, though it is not solely the province of the statistician.

I should like to touch briefly on one technical point. This is the question of the amount of experimentation that is required and the suggestion at the beginning of the paper that all experiments are sequential, implying a sort of decision function approach. This is made more explicit on p. 13, where it is stated: "This phase of the statistician's duties is likely to be influenced by research on decision functions and related topics during the next few years".

I just want to say a word of warning here because I myself am responsible for initiating work on the economic aspects of experimentation. This has been misinterpreted as suggesting that the new outlook on experimentation is that of decision functions; I think this can do a lot of harm. Scientific research has not primarily as its object the making of decisions, it is concerned with acquiring and broadening our knowledge of natural phenomena, increasing its detail and expanding it into wider fields. On the basis of this increased knowledge we may expect to develop new and improved techniques of all kinds and the perfection and application of such techniques often requires a somewhat empirical approach. It is only when one is concerned with the experimental aspects of this slightly limited empirical field—essentially the determination of working rules by experiments, which admittedly is what a lot of agricultural experimentation is—that the arguments I

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

have used to determine the optimum amount of experimentation have any relevance. Even here additional dividends may well accrue from further discoveries which emerge from a critical examination of the results.

Dr. Finney also brings out the point again when he says that the single decisive experiment is a somewhat rare phenomenon. It is true that as statisticians we are seldom concerned with a single decisive experiment, since if an experiment is decisive there is no need to consult the statistician about it. Consequently the statistician tends to find himself dealing with a disproportionate amount of rather indefinite experimentation which is mainly concerned with the determination of quantitative working rules. This does not mean that the decisive experiment does not occur in research.

Dr. S. C. PEARCE (in seconding the vote of thanks): Dr. Finney covers a wide field very successfully and with a great deal of wisdom. He details the many points in a research programme where a statistician can help, provided first, that he is consulted, and further, that he himself takes a wide enough view of his duties. The paper is well designed to spread enlightenment in both directions. I am in very good agreement with what Dr. Finney has said, and anything I add is by way of underlining and amplification.

The most important thing a statistician needs to learn from the experimenter is the *purpose* behind every trial that is proposed, and in my experience it is at this critical point that liaison most often breaks down. It is fairly easy to talk to a man and to learn the circumstances in which he has to work, so much land, so much plant material and so on, and to be told which treatments he would like to see introduced; but it is much harder to grasp why, as a scientist, he wants to do one experiment rather than another, what he believes himself to know already, and what he hopes to find out—in fact, to explore his mind as well as the circumstances. The *purpose*, once grasped, must dominate everything else, or we may find ourselves producing a quite brilliant design for the wrong trial, and I believe that statisticians do on occasion more harm than good for this very reason; they fall into the trap of being so absorbed in their own specialization that they forget the wider aspects of what they are doing. Thus, in designing a pruning trial on apples it is no doubt a good idea to make it factorial, but if two varieties are to be introduced we must know why. Are they expected to be so different that the best method for one is likely to ruin the other, or are they similar varieties introduced together for the sake of pollination? The distinction is important, for in the one case there must be enough replication to assess the pruning requirements of the two varieties separately; in the other, where information from one variety is likely to confirm that from the other, a lower degree of replication may suffice. Again, given a project for the comparison of ten spray chemicals, what really is the idea? Are all comparisons of equal value?—if so, the best plan might be to have, say, six randomized blocks with each chemical represented once in each. Or is the purpose to compare each of nine new chemicals with the present standard one?—if so, it might be better to triplicate the standard and to have five blocks each of twelve plots. If such simple trials call for an appreciation of purpose, how much more do complicated ones! One piece of advice that must be constantly reiterated is never to design a trial or analyse a body of data without a completely clear idea of what the experimenter wants to know, and the best way to find out is to ask the reason for the choice of proposed treatments.

I especially liked Dr. Finney's statement that there are two principles to be observed in all trials—the allocation of resources to best advantage, and the regarding of all experimentation as sequential. I think, however, that these two principles are sometimes rather difficult to see in perspective, especially when the trial is not of a purely empirical nature. When the purpose is, say, to assess the advantage in yield to be obtained by introducing a new variety, and the past evidence is summed up in a series of pilot trials, the position is fairly clear; but things are not so simple in dealing with an investigation of a more fundamental sort, where it is proposed to test whether or not someone's hypothesis provides a sound basis for thought. One difficulty is that the weight of existing evidence is much harder to assess, there being no measure of the goodness of an idea; another is the impossibility of assessing the value of right thinking about any subject, because there is no way of judging its consequences. What, for example, was the economic value of the discovery of Mendelian genetics? It is said that the discovery which before all others convinced man of his real place in the physical universe was that of the satellites of Jupiter, for it showed that the earth was not unique in an important respect; but it would have been very difficult at the end of the sixteenth century to have judged the value of a research project to discover if any of the heavenly bodies had satellites. In fact, the whole scheme would probably have been rejected out of hand as ridiculous and impious. I mention this point not because I think the work that has been done on the amount of experimentation is anything but useful for the problems that were in mind, but because I think we are increasingly going to be faced with a different kind

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of problem. Already the amount of knowledge in agriculture and horticulture is beyond the grasp of any single mind, and we must try to reduce it to principles, just as Kepler first reduced piles of data to his laws of planetary motion and Newton reduced these further to a single law of inverse squares. I do not suggest that biology is on the brink of a major development like that, but I do suggest that we must try to get beyond the *cataloguing* of phenomena and pass on to their *understanding*. To do this we must expect more trials that are not only steps in a sequence of experiments but also steps in a process of thought, and I believe that we shall have to think of them in different terms from those that suffice at present.

The vote of thanks was put to the meeting and carried unanimously.

Professor A. BRADFORD HILL: Having a proper respect for the duties of a speaker at a meeting of the Society I find myself embarrassed by my agreement with Dr. Finney's views on the statistician's multifactorial responsibilities. My only criticism would be that in some respects he is more diffident in his field than I would be in mine. For instance he feels it necessary to combat the notion that the statistician must confine his criticisms of a report to its statistical analysis or must refrain from commenting on the logical structure of an experiment. In the medical field we have long regarded it as our duty to be critical in whatsoever way the spirit moves us. To be able to do so effectively (and, indeed, safely) it is, however, fundamental for the statistician to acquire a very extensive knowledge of the field in which he has chosen to work. All statisticians have long taught that the experimenter, or observer, should not call us in at the end merely to carry out a post-mortem examination of the results. As F. H. K. Green observed, the statistician should be looked upon as an obstetrician rather than as a morbid anatomist. I doubt, however, whether the infant will prosper, or, indeed, even be brought to birth at all if we ourselves are entirely ignorant of obstetrics. We cannot merely "pass the buck" to our clinical colleagues; nor (and rightly) are they likely to listen to our advice or criticisms if we are obviously ignorant of medicine in general and, perhaps, of their particular field. As laymen we cannot learn the whole body of medicine nor, whether medically qualified or not, set ourselves up as rivals to Fellows of the Royal College of Physicians. But we *can* learn enough basically to talk good sense on any occasion when we are called upon to advise and be capable of taking up for special study in full medical detail every new project. I think this goes much further than Dr. Finney's plea (§ 3) that the experimenter "and his statistical colleagues should have sufficient understanding of one another's disciplines to avoid the errors that can result from imperfect communication of ideas". The ideal is that the experimenter should be his own statistician—that we should teach ourselves out of our job. That seems unlikely to occur. It follows that the statistician should not be a creature apart, merely occupied with techniques and tricks, but should get as near to the experiment and the experimenter as possible. I strongly support Dr. Finney's suggestion (§ 9) that he should normally be at a research institute in day-to-day contact with experiments and I agree that "he will not waste his time, bias his judgment, or harm himself in any other way if he occasionally soils his hands by assisting in routine field work". All such habits will add to the facility with which he talks the "language". *Without* that facility he clearly cannot effectively play a part in "assembling and evaluating the evidence" of published literature (§ 5). *With* it he should nullify Dr. Finney's dictum (§ 14) that "in general he will not himself be responsible for initiating the research". I see no reason why the statistician should not frequently be so responsible. I believe he should take every such opportunity and, indeed, in some instances carry out the research himself.

This fundamental need for an intimate knowledge of, and interest in, a particular field of work has a second implication, namely that, after a few youthful indiscretions, a statistician cannot lightly flit from one subject to another. Some persons, including statisticians, appear to think otherwise. Not only do they undertake to give advice wildly outside their own fields but quite often I get letters to say that the writer is tired of market research, quality control, or whatever it may be, and at the age of 48 is now ready to turn his hand (or machine handle) to the benefits of mankind in a worthwhile subject (such as medicine). I respect the ideals but my replies are not encouraging.

Dr. Finney writes that he has known "the value of a series of experiments to a more extensive summary to be completely destroyed by the discovery that in addition to some experiments reported in full numerical detail, there were others of the same series merely described as having failed to give significant results". That raises a point of some interest. I have known attacks made—not only by the harassed editors of overcrowded journals—upon the full publication of clinical trials in which the alleged remedy proved either valueless or no better than an established remedy. A mere *ex cathedra* statement is apparently held sufficient—or even silence. A negative result may be dull but often it is no less important than the positive; and in view of that importance it must, surely, be established by adequate publication of the evidence. But that is not at all generally accepted.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In medicine compared with any other field of experimentation there is one additional and over-riding difficulty—an ethical problem. When can one withhold a new (or old) treatment from any patient? Dr. Finney insists, for instance, upon the need for "untreated control plots" in addition to those treated in two different ways—since both treatments may be ineffective. In a clinical trial that may well be impossible. It was impossible, for instance, in a trial of cortisone in the treatment of rheumatoid arthritis. The hormone could be compared with the customary treatment by aspirin. "No" treatment was, and is almost invariably, quite out of the question. That is a problem of the greatest importance upon which I have spoken at length elsewhere and should not embark now. It is often overlooked by robust critics of the medical man's enthusiasm for adopting new remedies without adequate trial; when less robust the critics might look upon their own inclusion in the placebo group with some disfavour.

Mr. G. H. Freeman: More attention could possibly be paid to the importance at all stages in an experiment of collaboration between the statistician and the experimenter. It is desirable, as suggested in § 9, that the statistician should assist in field work, but more than desirable that he should have seen the land on which the experiment is to take place. All sorts of things that the experimenter knows so well that he does not bother to mention may influence the design, e.g. shading effect of hedges, slight dips indicating need for blocking one way rather than another, proximity of sources of infection. All these will be useful for the statistician to know, and provide a further argument for the location of statisticians at research institutes rather than elsewhere.

During the course of an experiment, or perhaps during the design, the statistician may be able to suggest the types of record that should be taken. Too often only one thing is recorded—usually total crop—while other records can usefully be taken that will provide more than subsidiary information. These include various measures of quality as well as category records on the growing crop and well repay the extra cost as the main cost is in the actual performance of an experiment while records are relatively cheap. This is perhaps the place where the statistician's help is least sought and most needed.

At the end of an experiment, when it is written up, the statistician also has a part to play. Few things look worse in a paper than undigested significant differences; indeed I would almost advocate a ban on the presentation of significant differences in print. The statistician may be able to help by translating his own technical jargon into English; if he cannot, it probably means he has not understood what he has been doing himself.

Mr. G. M. Clarke: Dr. Finney remarks that many experimental scientists have a knowledge of statistical principles and techniques. It is probably these workers who are happiest to have discussions with statistical colleagues, and not, unfortunately, those to whom such discussions could be of the greatest value. Such shyness, if I may so describe it, sometimes springs from reaction to University statistics courses, allegedly for scientists in general, but given by mathematics departments. A biologist, failing for some reason to benefit from such a theoretically biased course, becomes convinced he will never understand statistics, and so avoids it at all costs in his subsequent research. This, it must be admitted, is a more honest approach than saying: "I don't know what this statistician fellow is talking about, but I'll do it if it gets me the result I want". The likely result of lack of statistical thought, the laying down of a non-randomized, non-replicated experiment, is often defended by the cry: "If the differences aren't large enough to be obvious in the field, they won't be of any use to the practical farmers". This statement will be open to serious question, and at that point, as Dr. Finney reminds us, the words "demonstration plot" are thrown into the argument for good measure. Whether such non-statistical attitudes are tenable depends largely on the existing reputations of the experimenters involved; if a long-established laboratory worker suddenly decides he would like a field plot to "prove" some point of which he has convinced himself, the statistician is unlikely to hear about it until the "results" are published.

At the other extreme, there are experimenters who wish to take a number of sets of records, some seemingly irrelevant, and to perform as many statistical analyses as possible. We experience this in the horticultural world, not perhaps unexpectedly, since the number of long-term field experiments which one horticultural worker can sponsor in his working life is comparatively small, so that each must yield its maximum value. One aspect worth mention here is the practice of taking visual records or "eye-estimates" of such things as leaf-scorch caused by disease or nutritional deficiency. Such methods may give information useful at the design stage, a good example being to take rough assessments in the spring of blossom categories on a uniformly treated plantation of an apple variety which is to be used for an experiment on fruit thinning sprays. Visual records

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

110637

in their raw state cannot form variates for final analysis in an experiment, but I think there is scope, at least with some crops, for refining the method by serious attempts to remove observer bias, so as to bring it in line with a scoring system on, say, a 5-point scale. Attempts to quantify rough visual impressions often cause surprises. A colleague of mine recently took measurements of heights on several groups of nursery stock and found that the order of variability he felt he could see between the groups corresponded to the numerical order not of variances but rather of coefficients of variation. To him that was a more surprising result than it would appear to a statistician able to share the experimental experiences of colleagues with a wide range of interests. By this range of experience, a statistician should acquire in time a mental resilience which is quite a part of his general usefulness to his institution, and in saying this I support strongly Dr. Finney's comment (§ 9) in favour of a policy of providing resident statisticians wherever their services may often be required.

Another subject on which the statistician must exercise a restraining influence is in the volume of figures, the form of tables and so on, which his colleagues wish to include in their reports—often, he may feel, with a glow of satisfaction, because they attach more importance to them than he does. Words always mean more to an outside reader than the figures on which they are based, and it is not, I think, going too far to say that unless an experimenter can translate all his figures into words, he has not fully understood the implications of his results.

I would make one point of field design: the inclusion of untreated controls in entomological experiments is subject to theoretical difficulties, in that plots adjoining controls often have chances of being attacked to a different extent from plots which are surrounded by reasonably effective treatments; for this reason the "control" may well consist of an existing standard insecticide of known efficiency against which others will be tested.

Dr. Ruben: Dr. Finney has presented a powerful case for the closer integration of the statistician at all levels of an experimental programme and more especially in field experimentation. To the many convincing arguments which he has brought forward, may I add one more, although this is of a more narrowly professional character and likely to appeal, if at all, to the statistician rather than to the experimenter. In the hierarchy of the various branches of applied mathematics there is perhaps none in which the theoretical framework is more closely tied to and dependent on practice than statistical science. This would suggest that one can hardly expect real theoretical advance in this science which has not had its source in practical requirements and practical considerations. It would be interesting to examine the history of statistics over say the past fifty years from this point of view. I believe that such a history would emphasize fairly sharply that every major theoretical innovation in statistics, even if of a seemingly abstruse character, has arisen at one stage or at several stages removed from some practical problem. This point, though not new, is worth reiterating since there is a danger once it is forgotten that our subject may stagnate and that "mathematical statistics" (in my opinion, a dangerous misnomer) may produce a mass of mathematically imposing but essentially sterile trivia.

Dr. Finney has touched on the possibilities of sequential experimentation in branches of agriculture other than that of field experimentation and has advocated a wider though discriminating use of factorial designs. I would certainly associate myself with this recommendation. In this connexion, perhaps it may not be altogether without interest to mention some work done over the past few years in collaboration with Dr. K. Blaxter at the Hannah Dairy Research Institute on the efficiency of dairy cows (the results of which are due to be published shortly at various levels). Sir Ronald Fisher might be said to be the spiritual midwife of this work inasmuch as it was inspired by some remarks of his similar to those made in the issue of *Biometrics* referred to by Dr. Finney. Our study was undertaken in order to ascertain whether in an efficient scheme of herd improvement the efficiency with which cows utilized their food could be used to maximize use of land and resources. This problem falls naturally into two phases: (i) how can the efficiency of the individual cow with respect to the food intake be measured? (ii) given a method of measurement, what differences in milk yield and food requirements exist between cows when they operate at maximal efficiency, and are any such differences hereditable? But to date only the problem of measurement has been dealt with. Numerous interesting results have emerged from this work, but I shall confine myself here to one aspect which has a bearing on the design of dairy cow feeding trials and its possible implications on experimental design in other fields of experimentation.

Perhaps the most characteristic feature of the yield of milk in relation to food is that the magnitude of response of a cow to food depends on the sign of the change in food. More specifically, if a comparison is made between two planes of nutrition A and B the response, measured as a difference in yield, depends on the *order* in which treatment is given, even though the rations have been given long enough for an equilibrium (or an approximate equilibrium) of yield to have

R
519.905
J82J v.119
4257

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

been established on each ration. This effect is not to be confused with the secular decline sometimes (though by no means always) associated with lactating cows, and it leads directly to the phenomenon of irreversibility, i.e. a depression of the level of feed followed by a reinstatement does not result in a reinstatement of yield. The yield of the cow at any period of time is therefore a function not merely of the current feed intake nor yet of the simple aggregate of current food intake together with all past food intakes, but of the *pattern* of all past intakes. (If $Y(t)$ and $F(t)$ are the yield and food intake at time t, $Y(t)$ is a functional of $F(t)$.) We may use a sequential technique in such a set-up, but I believe, in view of the considerations just mentioned, that it cannot be usefully combined with a factorial approach, in which time is taken as one of the factors, if our aim is to predict the course of lactation corresponding to any prescribed food pattern. Consider, for example, 4 yields of a single cow corresponding to two planes of nutrition A and B both of which are given in each of two different periods. On a factorial approach, the 3 available comparisons from the 4 yields would be identified as 1 for rations on 1 degree of freedom and 2 for a period and period-ration interaction effect on 2 degrees of freedom which would normally be decomposed into their constituents. In point of fact, such a decomposition has no physiological basis and the interpretation of the 2 individual degrees of freedom is not at all clear. This would appear to have fatal consequences for the ingenious change-over designs which have been advocated for nutrition trials such as occur in dairy cow experimentation. These designs envisage the existence of a "main effect" due to current feeding and "carry-over effects" due to feeding in previous periods, the current yield being an additive function of the main and carry-over effects. The latter effects ignore the irreversibility phenomenon which contributes a large part to the rate of fall of yield during a sequence set of feeds; this rate of fall is an average effect determined, it is true, by the particular feeding sequence but not measured by the carry-over device. Such an approach, while it may give empirical results with a quite high degree of precision, fails to allow prediction of yield from a feeding sequence other than that employed to obtain the information. Another criticism which might be levelled at existing methods of dairy cow experimentation is that they tend to minimize experimental effects and produce bias, since the error variance in the analysis of variance tends to be very much smaller when real experimental differences are absent than when they are present. This disturbing effect is too deep-seated to be removed by recalibration, i.e. by a logarithmic or some other transformation.

These considerations may have a fairly wide field of applicability. For any input-output system in which the imposition of a stress affects its entire future behaviour (so that reciprocally the output at any instant is a function of the entire past history) a factorial approach, or indeed an approach based on experimental design in the narrow statistical sense, is liable to mislead. Such systems are historical and irreversible; they must occur frequently in biological and psychological investigations in which *time* is an essential element, as opposed perhaps to say mechanical and electrical investigations. The problem of the definition and measurement of a "treatment" as well as the interpretation of "error" is then a most difficult and elusive one. If we write $R_i = S + T_i$ where R_i and T_i are the response and treatment, measured on a given set of scales, for the ith period, and S is a parameter which is either an absolute constant or a predetermined function of time associated with the subject under investigation, then this statement is at best a tautology which serves to define the treatment in terms of the present condition of the subject and the latter's reaction. Thus in the dairy cow experiment the value of the food is as much a property of the animal as it is of the food itself. If, on the other hand, one writes $R_i = S + T_i + e_i$, where e_i is an "error" term, then the latter is in reality a complex parameter which serves to subsume the entire history of the subject as well as any possible experimental error and this is not vitiated if the model is rendered non-additive by including interaction terms on the right. In brief, such systems interact (using the word in its non-statistical sense) so closely with their environment and are so closely conditioned by their past history that (if I may be forgiven the metaphysical terms) cause and effect and subject and object cannot so to speak be profitably isolated, and a synthetic rather than an analytic approach is the more fruitful. This is perhaps rather reminiscent of the famous dictum of Heraclitus that you cannot step into the same stream twice.

The present line of thought leads me to suspect that Dr. Yates's remark (quoted by Dr. Finney on p. 10) that "all arguments which attempt to show that factorial design is ineffective or misleading when interaction exists contain their own refutation", while pointed, is not necessarily universally valid. For even if the "interaction" of period with treatment, on some scale of measurement, for a particular sequence of treatments in time is found to be statistically non-significant, the apparently additive model so found may not be valid as a basis of prediction for another distinct treatment sequence. For that matter neither may the modified model with interactions included, when these have been found to be significant, be valid for prediction. This is not likely to apply to cases where the end-product is sharply defined as in most field experimentation with annual

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

or seasonal crops, but rather to "sequential" crops (e.g. milk yield, perennial crops such as grass, etc.).

Nothing is further from my mind than an attack on the idea of factorial experimentation as such, still less on the design of statistical experiments. These have firmly established themselves as some of the most powerful weapons in the statistician's armoury. Heaviside once said of Fourier's work that it is the only genuinely amusing thing in pure mathematics; one might perhaps say of Sir Ronald Fisher's and Dr. Yates's work in factorial experimentation (and experimental design generally) that it is the only genuinely amusing thing in statistics. It has certainly given at least one statistician many hours of intellectual enjoyment. The aim of my remarks has been the modest one of attempting to show that in certain fields of inquiry and experimentation the factorial approach is inappropriate.

Mr. F. J. Anscombe: I should like to add my congratulations to Dr. Finney on this interesting paper.

It seems to me that in the last three or four years there have occurred two events of great importance and interest, likely to be regarded later as landmarks in the history of experimental design. The first is the short article by Dr. Yates in *Nature* on the amount of experimentation that is worth while in developmental work. Dr. Finney has been properly conscious of the economic side of the topics he has dealt with, but I am a little disappointed that he has not been able to say more about these things. The economic arguments used by Yates were rough-and-ready, and his article was very brief. I think it would be enormously worth while for a much fuller economic investigation to be carried out. I should have thought, in fact, that the Ministry of Agriculture would attach great importance to a careful economic study of agricultural field experiments; and that some industrial concerns would have a similar interest in industrial experiments. I should like to ask Dr. Finney whether he knows of any follow-up of Yates's work, and, if nothing has been done about it, whether he can suggest any reason for this. (I am referring to an *economic* follow-up, and not to any development of statistical theory like that of Grundy, Healy and Rees.)

The other event I was referring to is very recent, in fact not yet fully published, namely, the work of Dr. G. E. P. Box and his colleagues on factorial designs. This gives us a new insight into the nature of factorial experimentation. It is appropriate to mention this work now, because of Dr. Finney's discussion of factorial experiments in the paper.

Dr. Vajda: It has been said to-day, as it has been said before, that it is important to bring the statistician into the procedure when the experiment starts. I fully subscribe to this but must point out that very often it does not happen. I do not see any point in telling our students and young colleagues that they should be cross if they are not brought into the experiment at an early stage; it is more important to tell them what happens if they are brought in too late.

My work is in the Admiralty and I cannot always blame naval officers for not knowing when they start experimenting with torpedoes and so on that some statistical problem might be involved. I think it clever of them to realise that it might be interesting to bring in a statistician and ask relevant questions, which in my opinion must be answered.

May I plead that we should teach those whom we are called upon to teach what to do when information is lacking. Our literature contains some aspects of this problem and I think much more could be done. I should like to see an examination question set: "You are brought into the proceedings at a very late stage, can you still do something about it, and if so, what?" I mention also the economic aspect which Dr. Finney insists is important. It might be worth while investigating what one loses by being brought too late into an experiment. I suggest that there are two aspects of this. You may either lose something by having to do harder work, or what you can do at a later stage is less useful. If this could be expressed in some measure, such as money, which people understand, we might start some propaganda among the people who call us in too late.

Mr. J. A. Nelder: We must be grateful to Dr. Finney for amplifying and clarifying many general problems which statisticians are continually meeting in the course of their work. I have kept my points as general as possible, but I have seen them mostly from the standpoint of horticultural research.

The role of a statistician in selecting treatments is very tricky. I agree with Dr. Finney that a statistician can sometimes improve the structure of the experiment so that it gives proportionately more information by a change in the treatments. I would suggest, though, that a good experimenter will know just what he wants to do and will have adjusted the treatments to fit his requirements. However, one sometimes finds a situation where someone seems to have said "here are

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

some resources, let's do an experiment", for experimenting, as we all know, is, in Mr. Potter's terminology, an "O.K.-activity". A few general factors are proposed and it is left to the statistician to design the experiment on this basis. There is a temptation here for the statistician to think up an appallingly complex experiment designed to answer by next year all the questions anyone might ever ask about the crop. He should, I think, oppose doing the experiment at all. Briefly, if he finds he is having the chief say in the choice of the treatment structure, this may be evidence for the incompetence of the experimenter, and so a reason for opposing the whole experiment.

I would support most strongly Dr. Finney's emphasis on the economic aspect of treatments in applied research. It is surprising how rarely this seems to be brought into the discussion of results. In attempting to assess the benefit accruing from a new treatment it is most important not to forget the length of time it takes for a new practice to become generally used by the farming community. It is not enough merely to multiply the increase in crop caused by the treatment by the acreage of the crop in assessing the value of the treatment.

The question of plot size is extremely complicated. I notice that most of the calculated optimum plot sizes published are much smaller than those normally used. This is partly because a minimum plot size is frequently dictated by the terms of the experiment, but experimenters do genuinely fear that a large number of very small plots are more likely to lead to a disproportionate number of errors in application of treatments and recording. My own decision rule for vegetable crops is "use as small a plot as the treatments allow", and I should be interested to hear Dr. Finney's views on that. I have not convinced some experimenters of the optimum character of this decision rule.

I am, in general, very sympathetic to Dr. Finney's advocacy of multifactorial experiments. Much depends on the technical ability of the experimenter to bring the experiment through without a hitch. Where the "clot-factor" is prominent, fractional replicates and the like are dangerous, for the "clot-factor" is never orthogonal to the rest of the factors. There is a danger of misinterpretation, too, where certain treatments alter error variances; in fact the designs suffer from a certain lack of "robustness", to use Box and Anderson's terminology.

May I make one final point on replication. The formula giving the variance of the mean as σ^2/n should not be extrapolated too far, for I am sure that often σ^2 is an increasing function of n. This is especially true where skilled labour is short. Again a decision rule that covers many situations is "do as much as you can do well".

Dr. SAMPFORD: I should like to comment on three points in Dr. Finney's paper, and to reinforce them with awful warnings from my own experience.

Firstly, a decision should be taken at the outset on what measurements are to be made, and any limitations imposed by the treatments on these measurements should be borne in mind when choosing the design. I have on several occasions recommended a latin square as being a suitable design for an experiment, only to find when the results came in that although the latin square *was* suitable for the measurement of primary interest, the experimenter had also measured various secondary characters that for some reason could not be observed on plots receiving certain treatments. In consequence, not only was the analysis of these secondary characters complicated by the latin square restraints, but the number of error degrees of freedom sometimes became dangerously small: in some, if not all, of these experiments the choice, instead of a latin square, of a randomized block design occupying the same number of plots would have avoided these difficulties without increasing too seriously the error variance of the primary measurement.

Secondly, the statistician should attempt, with the help of the experimenter, to visualize all possible effects of the treatment, and to choose a design the desirable features of which will not be vitiated by the nature of the responses, even if it is not at first sight the most convenient. I was once asked to design an experiment on a standing tree crop, in which the treatments were to be six levels of fertilizer, and two types of cultivation. The available stand was very heterogenous in growth so I proposed initially a latin square design for the fertilizers, with each plot split for cultivation, as this was the factor of primary interest. As it afterwards turned out, this design would have been satisfactory from a statistical viewpoint: however, we were forced to abandon it because of the difficulty of cultivating small areas, and the design actually used was a half-plaid square, in which the cultivation treatments were applied to columns of the fertilizer square. Unfortunately, the two cultivation treatments produced such startlingly different effects, and such very different residual variances, that we were obliged to analyse the two halves of the experiment separately, as two semi-latin squares, a procedure that resulted, once more, in a lot of unnecessary work, and dangerously few error degrees of freedom.

My last example is put forward to reinforce the suggestion that the statistician should keep track of everything that goes on during an experiment. I have recently been concerned with a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

grassland trial carried out at a number of different sites: details of treatments, sowing, and harvesting were carefully standardized over all sites, with one exception. One area was grazed during one year of the trial : this was not regarded as a treatment, but merely as a routine to keep the sward in good condition for treatments to be applied in the following year. For this reason, presumably, no specification was made of the number of animals to be used or of the grazing rotation to be adopted, with the result that these features showed considerable differences between sites. This experiment was not, in fact, designed by me, but I think it quite likely that I too would have overlooked this detail: Dr. Finney's ideal statistician would not have done so, and rightly, for the effects of such an oversight might be serious, although in this case, fortunately, they do not seem to have been so.

Mr. SKELLAM: Most speakers have taken up the point that there should be collaboration between the statistician and the experimenter. It is easy perhaps to establish such collaboration, but it is far from certain that the collaboration will be fruitful, for the simple reason that the two people often speak a different language. If the experimenters' aims are comparatively simple, if he just wishes to know whether a particular combination of treatments is best, the statistician perhaps has no difficulty, but if the experimenter is concerned with the way that nature works, with the processes that are going on, it is important that the hypotheses against which the data are compared should be hypotheses which are scientifically realistic, not some linear combination of the variables because they can be easily treated mathematically. I think it important that the statistician should try very hard, shall I say, to translate the ideas in the minds of the experimenters into some kind of mathematical form, not necessarily the mathematical form which is most easily manageable, but a mathematical form which the scientist would approve.

The CHAIRMAN: I feel impelled to make a few observations on this subject apropos one or two points mentioned in the discussion. Dr. Finney's paper is directed largely to field experiments in the sense of agricultural field experiments. The discussion has taken us a little further afield than that, and I think it should be made plain that a good deal of the experimental design which is useful in field experiments in Dr. Finney's sense is not necessarily applicable in the experimental design which we have to use elsewhere. I have in mind particularly the attempts which we have been making to introduce experimental designs into experimental inquiries in sociology, social psychology and economics. Although I do not dissent from anything Dr. Finney has said in the paper in its context I think it might be well worth while emphasizing that some of these statements might be open to objection if they were carried into other fields; for example, the statement that random sampling is almost *sine qua non* in an experimental design cannot be applied in a good many social inquiries. I think of the famous inquiry of Dr. Doll and Professor Bradford Hill on cancer and smoking. In this instance they were practically compelled to use matching techniques, a subject on which we might well find it useful to know a great deal more. How far their inquiry can be described as an experiment is perhaps arguable, but if it was not an experiment it was an experimental inquiry.

The other point in connection with agricultural design with which we are finding difficulty is in the degree of elaboration. A multi-factorial experiment is a very fine thing if you can carry it out, but it very often happens, especially in sociological work, that a piece of the design turns out to be missing. This is not merely a question of the missing plot; you may have a whole missing treatment, as when a participating organization may find that it cannot do field work at a particular time. The more elaborate designs have a certain filigree quality which makes them particularly susceptible to damage. What we want in this particular field is not a robust test but a robust design, a design which can stand up to a great deal of hard treatment in the field. Such considerations are apt to be overlooked if we turn our attention solely to those fields where we can carry out what we intend to do.

The following contribution was received in writing after the meeting:

Mr. C. P. COX: I particularly welcomed this definitive presentation of the functions of the statistician in experimentation since I gave a colloquium on the subject at my Research Institute five years ago. At that time I had the benefit of Dr. Finney's remarks on my paper and I am pleased to find that we are still in agreement—particularly since I have recently been asked to repeat my colloquium.

Many authors have considered general aspects of the question, but it is striking that most expositions on the subject have emanated from statisticians, not experimenters, and I suspect that most practising statisticians will be aware of the general issues involved. The real problem,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

however, remains that of convincing experimenters themselves of the value of these widening activities of statisticians in research.

I have been fortunate in achieving excellent collaboration, including participation in experimental work, with a number of colleagues, but although the example is spreading, this "contagious" process is slow. One factor is that, of historical necessity, heads of experimental departments have often achieved their puissant eminence without the benefit of statistical methods; when such leaders become encouragingly aware of the quantitative approach progress is more than individual.

I have found that collaboration in a wider role has associated drawbacks. Dr. Finney has discussed an ideal relationship between the statistician and the experimenter; in practice, however, it is often the one statistician and up to twenty experimenters, a somewhat extending situation for the former. Further, the introduction of more efficient designs usually leads to an increased amount of experimentation and further attenuation of the statistician's resources. Until, therefore, these are commensurate with the wider responsibilities rightly being advocated, I think that the increased value of statistics in research can best be achieved by addresses to experimenters, working from the top of the research chain downwards.

Dr. FINNEY: I am in agreement with almost all that contributors to the discussion have said, and to most of these my reply will be: "You have put your emphasis a little differently from my own because you have a different background, but had I thought of all these points I would gladly have incorporated them in my paper—so making it twice as long".

The Chairman's remarks give me the opportunity of emphasizing that I have been speaking about agriculture. I do not wish to dogmatize about the applicability of the principles I have discussed in medical or sociological research; I am anxious only that their relevance should be examined. I suggest, however, that the type of investigation he has mentioned is a sample survey and not an experiment in the sense generally accepted by biological and physical scientists. Sample surveys have their own problems of planning and design, related to but not identical with those of experiments in which the allocation of treatments can be controlled. In particular, random selection of units for observations is not the same thing as randomization as a determinant of treatment.

Dr. Finney subsequently replied in writing as follows:

I was glad to hear Dr. Yates's insistence that experimentation is not only a matter of decision functions, and I hope that nothing in my paper seemed to imply a contrary belief. It is well to remember that a good experiment can add to our knowledge much more than "yes" or "no" on half-a-dozen direct questions, and often more than was in our minds when we planned it; in agricultural research especially, I suspect that we still do not extract from experimental results all the information that they contain. Nevertheless, many decisions are taken by scientists other than those relating to the empirical determination of working rules—for example, decisions as to whether a particular line of research in pure science should be continued, whether the time has come for trying the effect of a change in temperature in a process under investigation, or even of whether or not a paper summarizing a piece of research should be prepared for publication. I do not say that these are fit subjects for the formal theory of decision functions, but I think it would be legitimate and perhaps even useful to inquire into the relationship between the evidence and the decision taken. Dr. Pearce has pointed out the difficulty of formulating a scale of values in circumstances such that an economic scale is irrelevant. He is obviously right, yet every research worker whose position allows him any control over the planning of his work is frequently forced to make at least a rank-ordering of values. Moreover, this ordering appears to be not entirely subjective. We have no measure of the goodness of an idea, yet research workers experienced in a particular field will often agree on the goodness and importance of an idea; there is scope for both statisticians and psychologists in studying their agreements and disagreements.

Dr. Ruben will not expect me to comment in detail on the exceptionally interesting experimental situation that he has outlined. I shall await impatiently the publication of a piece of research that evidently conforms to Dr. Pearce's and Mr. Skellam's ideals in its attempts to understand phenomena, and not merely to catalogue them. On one small point he is perhaps a little unfair, for, in his criticism of a remark by Dr. Yates that I quoted, he interprets *interaction* in a sense that includes the type of complexity he encountered with dairy cattle but restricts *factorial design* to the description of a particular category of familiar designs instead of allowing it to include all designed experiments in which several factors must be studied. In the broader sense, his type of investigation could also be termed factorial.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Dr. Yates may be able to give Mr. Anscombe a different answer, but I do not know of any direct follow-up of the work on the economic amount of experimentation. Series of experiments that Dr. Yates and I helped to plan in India gained something from this important idea, but not to the extent of having their number exactly determined by Yates's formula! Fuller information on the costs of experimentation is needed, and I would like to see attempts made to compile this. With what Mr. Anscombe has said about Dr. Box's work I agree completely; I omitted to mention this development of factorial design solely because field experimentation at present seems one of its least promising applications.

Of course Dr. Vajda is right in saying that statisticians must be prepared to advise on the interpretation of experiments that they have not planned, but I was concerned primarily to emphasize the ideal of full consultation. His suggestion for measuring the loss from inadequate statistical help is interesting and worth further examination. I am less willing to agree that questions about torpedoes "must be answered", for this seems to imply a scale of values that I cannot accept.

I want to thank other speakers for many interesting comments, but I think I need not refer to these in detail. Professor Bradford Hill, in particular, expresses himself more strongly than I had dared to do; I can but hope that, when I have his length of experience, I shall also have his courage and possibly deserve equal respect for my views.

As a result of the ballot taken during the meeting the candidates named below were elected Fellows of the Society:

Meher Abbas.
Abdel Azim Anis.
Suliman Taha Ayoub.
Francis William Bacon.
Brian Oliver Bartlett.
James Edward Bessell.
Raymond Charles Bond.
Norman Nathan Bronstein.
Victor Edward Bruce.
Lesley Mary Chuck.
Arthur Wellesley Stewart Clark.
Charles Alan Cotton.
Robert Nicholas Curnow.
Leonard Ernest Davison.
John Cecil Dickins.
Geoffrey Charles Doepel.
Edward Frans Drion.
Hunter Duff.
Ronald Victor Dulley.
John Henry Farnfield.
James Edward Sullivan Figgins.
Alan Frederick Findlay.
Allan Fletcher.
Leonard James Flowerdew.
Kuno Ruben Gabriel.
Jack James Grimshaw.
Sidney Martin Helps.
Joan Hockridge.
Margaret Eleanor Holgate.
Anthony Victor Brian Horsfall.
George Peter Hyett.
David Noel Isserlis Ibbetson.
Frederick Henry Jackson.
John Stanley Jackson.
John Norman Richard Jeffers.
Edward Leslie Jenkins.
Alewyn Jones.
Arthur Wolfe Joseph.
Bernard D. Karpinos.
Zbigniew Alexander Lomnicki.
Jack Gentry McGlennon.
Philip Martin Matthews.
Dennis Michael Milstone.
Kazimierz Morawski.
Norman Ernest Newton.
Malik Mahmoud el Nomrossy.
John Michael Playfer.
Graham Frederick Reed.
Edmund Augustus Richards.
Erica Royston.
David Bernard Shine.
Joan Short.
Janusz K. de Somogyi.
Mervyn Stone.
Frederick William Thomerson.
Gordon Blair Trasler.
Peter Kaineng Tsao.
Swaminathan Vamathevan.
John Ernest Weaver.
Denys Stuart Milroy Webster.
John Francis Woods.
Stanislaw Krystyn Zaremba.

Corporate Representatives

Arthur Graham Francis Farquhar, *representing* The Central Electricity Authority.
Gordon Norton, *representing* Orbit (Operational Research) Ltd.
Arthur Asher Shenfield, *representing* The Federation of British Industries.
Anthony George Starr, *representing* The Ministry of Supply.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Causal Inference from Observational Data

A Review of Ends and Means

By Herman Wold
(*Uppsala, Sweden*)

[Read before the Royal Statistical Society, Wednesday, December 21st, 1955, the President, Professor E. S. Pearson, C.B.E. in the Chair]

Contents

PAGE

Introduction and Summary 28

1. A Fourfold Map of Ends and Means in Applied Statistics . . . 29
   - 1.1 Description *v.* explanation, a dualism in ends 29
   - 1.2. Observational *v.* experimental data, a dualism in means . . . 30
   - 1.3. A fourfold map of ends and means 31
   - 1.4. Three illustrations from the SW sector 32

2. A Review of Explanatory Approaches, with Particular Regard to the Treatment of Experimental *v.* Observational Data 34
   - 2.1. The specification of hypotheses 35
   - 2.2. Estimation of parameters and other elements of causal inference . . 38
   - 2.3. Hypothesis testing 39

3. The Explanatory Analysis of Observational Data: Some Typical Situations 41
   - 3.1. Causal patterns 41
   - 3.2. Regression analysis 42
   - 3.3. Model building on the basis of elementary assumptions . . . 47
   - 3.4. Joint treatment of causal and stochastic assumptions 49

List of References 49

Introduction and Summary

Intensive development of statistical methods in the last three or four decades has created broad areas of applied statistics whose techniques possess a high degree of exactness and efficiency.

On the one hand, procedures for sampling surveys supply rational techniques for mass observation and for the assessment of frequencies, averages, and other elements of collective description. On the other hand, the rules for the design and analysis of experiments embrace both descriptive and explanatory problems, with emphasis on explanation. Modern methods thus treat three categories of problems: (1) description on the basis of observational (non-experimental) data, and (2, 3) description and explanation by the use of experimental data.

A fourth broad area remains—explanation on the basis of observational data. In this field, which embraces among other things a large proportion of social science research, progress has been less systematic and spectacular. Current text-books reveal the step-child treatment which this sort of problem has received from professional statisticians. Nevertheless a small cadre of statisticians has made basic contributions in this direction, and these developments may be expected to bring the fourth area more into balance with other fields of applied statistics.*

The purpose of this paper is to examine the differences in explanatory procedures when dealing with the two types of data: experimental and observational. After introductory comments in

* In particular I am thinking of the situation in the social and economic sciences. The growing interest in causal models and their study by statistical method is reflected in Unesco (1954), Lazarsfeld, ed. (1954), Kempthorne *et al.*, eds. (1954). The pioneers include Wright (1934) and Tinbergen (1939).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

section 1 on the areas mentioned, section 2 continues with causal inference. It is argued that the theoretical model on which the statistical analysis is based can be specified with less exactness and detail when passing from experimental to observational data, and that this leads to a partial re-orientation of the statistical techniques. The shifts of emphasis are examined with regard to specification of hypotheses, estimation procedures, and hypothesis testing. The conclusion reached is that with observational data the statistical analysis becomes more dependent on (*a*) co-ordination with subject-matter theory, (*b*) large-sample methods, and (*c*) checks and tests against other evidence. In section 3 this conclusion is illustrated with reference to four types of approach. The illustrations have been selected so as to bring out the differences in the causal analysis with regard to complexity and aspiration level.

This brief essay treats a large topic and, as my experience and outlook are of necessity limited, the aim is to stimulate interest rather than to stress my own views. Such novelty as the paper may possess lies in the arrangement of material and the integrated picture of statistical method.* The examples are simple, many being drawn from my own work on demand analysis, a branch of applied statistics displaying typical problems in the causal analysis of observational data.

1. A Fourfold Map of Ends and Means in Applied Statistics†

1.1. *Description versus Explanation, a Dualism in Ends*

A descriptive approach answers such questions as "How?" or "How much?". An explanatory approach treats the question "Why?", and in some form or other the answer involves an element of causal inference.

A few examples: in agricultural research, the assessment of yield per acre is description, while an investigation of the yield in its dependence on manure falls under the heading of explanation. In medicine we may, on the one hand, refer to the weekly official reports on epidemic disease (description), and on the other, to clinical data on the curative effect of some new antibiotic on a specified disease (explanation). In industrial statistics we may cite the control chart for a certain element of operation along the assembly line (description), and a project of operations analysis which serves to improve the production process by removing the disturbing influence of some nuisance factor (explanation).

A frequent situation is that description serves to maintain some *modus vivendi* (the control of an established production process, the tolerance of a limited number of epidemic cases), whereas explanation serves the purpose of *reform* (raising the agricultural yield, reducing the mortality rates, improving a production process). In other words, description is employed as an aid in the human *adjustment* to given conditions, while explanation is a vehicle for *ascendancy* over the environment.‡

While a statistical investigation may have no purpose other than description, an explanatory investigation always involves an initial phase of description, that is, of orientation and of fact collecting to provide material for the causal analysis. From this point of view a complete research project involves three phases:

(i) Description; including primary measurements, tabular and graphic presentation of facts and the assessment of collective characteristics.

(ii) Explanation; including the specification of causal hypotheses and their testing against the empirical data.

(iii) Application. This final phase includes some element of prediction; that is, the inference drawn from the observed sample is extended to some other sample, different in time or space or otherwise.

* The paper was planned in consultation with Dr. R. Porter, Fort Walton, Fla., and G. Eklund and K. Medin, Associate Professors at the Univ. Inst. of Statistics, Uppsala. The first draft of the paper has received constructive comments from these colleagues and from Dr. C. A. Anderson, Univ. of Kentucky; Dr. J. G. Bryan and Dr. E. Robinson, Massachusetts Inst. of Technology; Dr. R. Strotz, Chicago; Dr. J. Tukey and Dr. M. Wilk, Princeton; and Dr. P. Whittle, Wellington, New Zealand. To them all I extend my heartfelt thanks.

† This introductory section is mainly based on Wold (1954). For the place of causality in the general theory of knowledge, cf. Braithwaite (1953), especially p. 311.

‡ For this last interpretation I am indebted to Dr. R. Porter.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

1.2. *Observational versus Experimental Data, a Dualism in Means*

Without comment we give as two examples of description: a set of measurements on the specific gravity of chromium (experimental data), and the divorce rate percentage of marriages (observational data).

As regards explanatory analysis, we shall briefly recount a classic in medicine, J. Dietl's experiment on pneumonia (1849). A hundred years ago medicine was still dominated by bleeding and emetics, treatments that had been in general use for all sorts of diseases from mediaeval and earlier times. Dietl made his experiment on three groups of pneumonia patients, 380 cases in all. In accordance with the traditional prescriptions two groups were treated with leeches and tartar emetics, respectively, while (because Dietl was courageous enough to put the established authorities to a test), the third group was given so-called expectative treatment: the patient was not treated at all but placed in bed and given light and nourishing food. The resulting mortality was 20·4 per cent. for the group treated with leeches, 20·7 per cent. for the group treated with emetics, and 7·4 per cent. for the control group. The cures thus were twice as dangerous as the disease itself! Dietl's report was one of the first agents in tearing down the citadel of mediaeval medicine.

The notion of experiment is somewhat fluid, but whether taken in a broad or narrow sense it involves some degree of planning and control from the side of the experimenter. We shall take here the term in a narrow sense which, when a distinct term is required, will be referred to as *controlled experiment*, and which we specify by requiring the following conditions to be fulfilled, more or less rigorously (see Fisher, 1935, esp. Sections 20–21).

(i) The replications of the experiment are made under similar conditions (so as to yield an internal measure of uncontrolled variation).
(ii) The replications are mutually independent.
For explanatory experiments the following condition is added:
(iii) The uncontrolled variation in the replications is subjected to randomization in the sense of R. A. Fisher.

The causal hypothesis of an explanatory experiment usually takes the form of a relationship in which the variable under investigation, called the effect variable, is expressed as a function of causal variables or factors. Some of the causal factors are controlled in the sense of being subjected to systematic variation, others are uncontrolled. The randomization under (iii) serves to make the controlled variables independent of the uncontrolled factors. Considering, for example, the case of linear regression, let

$$\eta = \alpha + \beta x + \zeta(\varepsilon^{(1)}, \ \ldots \ , \varepsilon^{(s)}); \quad y = a + bx + z \qquad . \qquad . \quad (1\ a\text{–}b)$$

be the theoretical and least squares regressions; the randomization makes the controlled variable x independent of the uncontrolled factors $\varepsilon^{(i)}$, and if the residual ζ does not depend on x it follows under general conditions that b will be a consistent estimate of β.*

Observational or non-experimental data are those in which one or more of conditions (i)–(iii) are violated. The empirical data are mainly experimental in some sciences, as physics and chemistry, and mainly observational in others, as astronomy, meteorology, economics and the social sciences; but most frequently perhaps the data are drawn both from experimental and observational sources, as in psychology, genetics and medicine. This is to some extent a question of convention, because conditions (i)–(iii) are somewhat vague, and the empirical data will therefore often have features in common with both experiment and observation. Agricultural field experiments are a case in point since the soil variations are not in complete agreement with (i) and (ii).

Assumptions (i)–(ii) are usually adopted in the strict sense of probability theory, the uncontrolled variation being assumed to form a set of independent variables with the same distribution function. The techniques for industrial quality control are a typical case in point. These require that the two assumptions be fulfilled by the control chart of the operation under investigation, and the operation is then said to be under *statistical control* (Shewhart, 1931; Hald, 1952, Ch. 13.2). In many time-series data assumption (ii) is violated by the presence of auto-correlation. This is one reason why times-series data will often come under the heading of observational data.

* For a more detailed treatment of the situation, including cases when ζ depends on x, see Wold-Juréen (1952), Exercise IV, 15.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Another is that if there are time trends in the uncontrolled variation their correlation with the time variable cannot be removed by randomization, for time is a factor which cannot be subjected to systematic variation.

1.3. *A Fourfold Map of Ends and Means*

The double dichotomy in ends and means briefly commented upon in section 1.1–2 gives us a fourfold map of statistical methods, as shown overleaf. The compass notations serve for easy reference. The classification of methods thus obtained is sufficiently clear-cut for our purpose. The methods of the SW sector being our topic, we shall now prepare the discussion in sections 2–3 by some comments and illustrations.

(i) Description involves simpler problems than explanation, and explanatory problems are easier to solve with the use of experiments than without. Hence the methods required are on the whole less penetrating in the NW sector of the map than in the SE, and the SE sector in its turn is less difficult than the SW. This situation throws some light on the broad lines of development of statistical method. Up to the beginning of this century statistics was by and large synonymous with the NW sector. Approaches into the other sectors were in the picture at an early stage; we note the construction of mortality tables (SW), Gauss's methods for the treatment of observational error (NE and SW), and causal experiments like that carried out by Dietl (SE). The epoch-making contributions of R. A. Fisher brought a powerful expansion in the direction of experimental methods, with emphasis on the SE sector. The exact and efficient methods developed by Fisher and his followers have led to great triumphs in biological sciences, medicine, engineering and other fields where research can be based on controlled experiment, and their range of fruitful application is still expanding. Speaking broadly, classic and modern statistics focus on the NW and SE sectors of our map, respectively. This suggests that the next phase of development will centre on the SW sector, and work in this direction is well in progress, as already noted.

(ii) In the SW sector the statistical methods cannot profitably be set forth in a general formal manner; they split up into a number of specialized topics, such as demography, econometrics, etc. (Wold, 1954, p. 170). This is because in the absence of experiments the statistical analysis has to be closely coordinated with subject-matter theory both in specifying the causal hypotheses and in testing them against other sources of knowledge. A crucial feature is randomization, which in experimental situations reduces the disturbing effect of uncontrolled variation. This device not being available in observational situations, it is a pertinent problem to what extent the disturbance factors should be taken into explicit account, and at this point strong reliance must fall upon subject-matter arguments.

(iii) As the lines of demarcation in the map are fluid, special attention is called to two zones: (a) between the SE and the SW sectors, (b) between the NW and the SW sectors. Both these zones are important research fronts where the exact methods of the NW and SE sectors are being generalized and adapted so as to cope with some of the problems of the amorphous SE sector.

An important array of intermediate cases between the NW and SW sectors stems from differential statistics: death-rates by sex, rural versus urban marriage rates, etc. An element of explanation lies in the differential factor; since, however, the dependent variable is usually influenced by additional factors, the explanation will only extend beyond the sample in a qualitative and unoperational manner. Emphasis of differential statistics thus lies on the description of the particular sample. Throughout the history of statistics applied work in the social sciences has in this way hovered between description and explanation. In favourable circumstances differential statistics have yielded results that were explanatory in an operational sense and hence definitely belong to the SW sector The earliest and best example is the computation of death rates by age and the ensuing explanation of mortality as a function of age.

Correlation is another intermediate case. If X, Y are two variables given by observational data, let

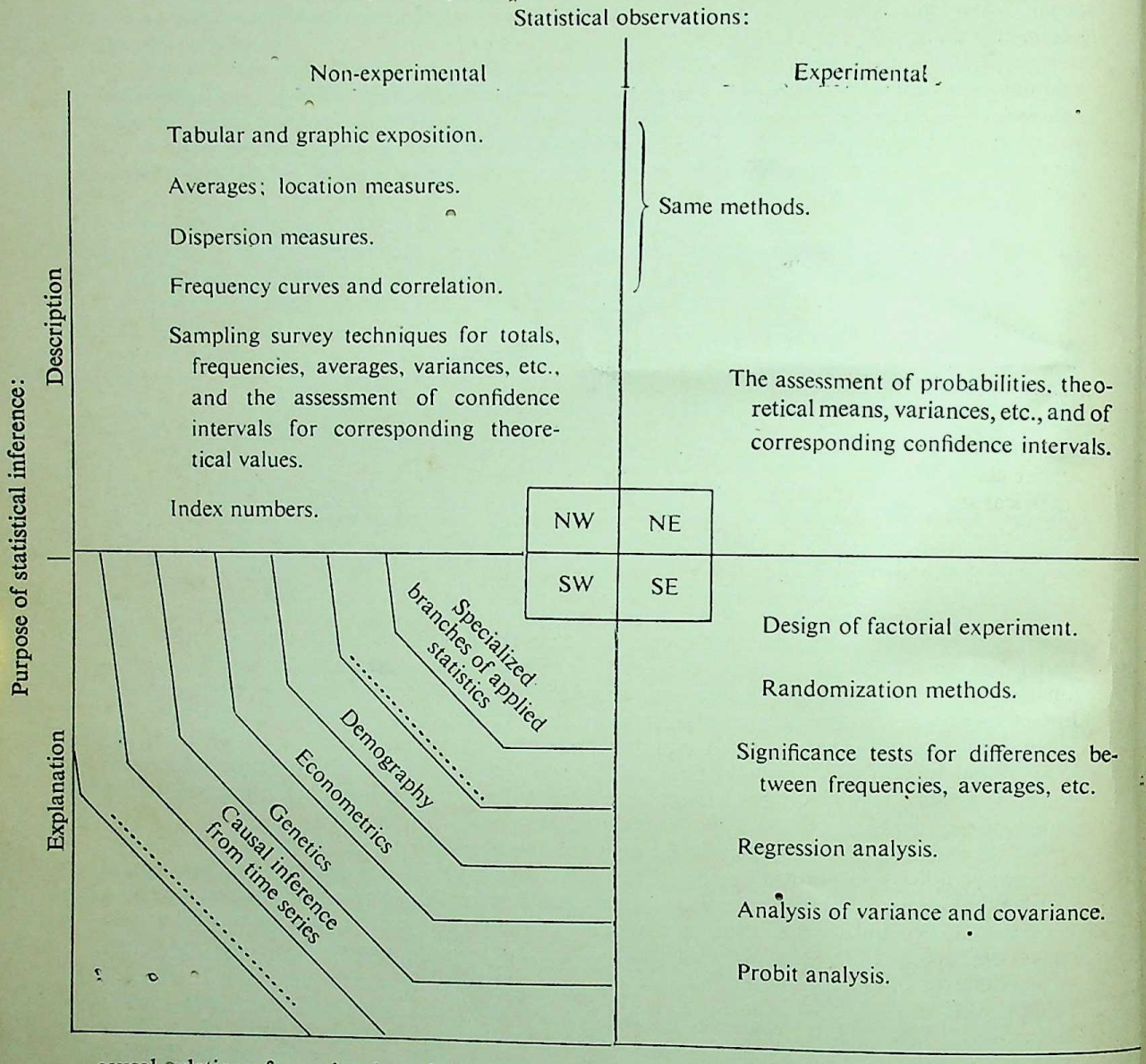
$$X \longrightarrow Y \qquad X \longleftarrow Y \qquad X \longleftrightarrow Y \qquad X \text{———} Y \qquad . \quad . \quad (2\text{ a–d})$$

indicate the hypotheses: (a) Y is causally dependent upon X, (b) X is dependent upon Y, (c) X, Y are intercorrelated without any causal explanation being specified, and (d) X and Y are causally independent. If the causal dependence is specified (cases a and b) we assign the case to the SW sector; otherwise (case c) to the NW sector. Regression methods, when used to extract

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Fourfold Map of Ends and Means in Applied Statistics



causal relations from the data, belong to the SW sector. Correlation coefficients and other measures of association fall to the NW sector.

1.4. *Three Illustrations from the SW Sector*

It need hardly be stressed that the SW sector is large and important, and its problems are of varying intricacy and difficulty. Innumerable examples could be cited where research has successfully coped with minor aspects but where we are far from having a firm operational grasp of the central factors. We shall return to this problem. One may refer to the notorious difficulties when dealing with clinical data in medicine compared with the treatment of laboratory data. (I think of the selective factors that prevent co-ordination of experience from different clinics.) The situation is similar in psychology, education, sociology, etc.; for instance, analysis of causal complexes underlying juvenile delinquency, alcoholism, or divorce. We proceed to discuss three illustrations in more detail.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(i) *Mortality tables.* The dependence of human mortality upon age is so obvious that it requires an effort to realize that this is in principle a causal explanation as hypothetical as any other. Some time after Halley's mortality table was constructed in 1693, such tables were used in life insurance calculations. The success of the business use demonstrated that mortality could be explained as a function of age, and that the inference from one set of data could be extended with sufficient accuracy to other populations. Two general features in this illustration are important: (a) observational data are used for the numerical assessment of some element of causal inference, (b) the application of the results involves a prediction by which the inference from the observed sample is extended to other samples. Among the many refinements we note Gompertz-Makeham's formula for functional representation of the causal dependence,

$$\mu_x = a + b \, . \, c^x$$

where μ_x is the instantaneous death rate, x is age, and a, b, c are parameters.

(ii) *Demand analysis.* The conception of a demand function was introduced into economic theory by A. Cournot (1838). The representation of demand as a function of price enters into the theoretical model by which he explained the stability of price, together with the balance of production and consumption in a freely competitive market. The same demand function is the basic tool in his theory of monopolistic price. Jevons (1871) stressed the basic role of the demand function and also emphasized the importance of obtaining empirical knowledge concerning demand functions. However, demand functions remained theoretical until the first estimate of a price elasticity was made by R. Benini (1907).* After the pioneer work of H. L. Moore (1919–25) empirical studies multiplied, and were followed in the 1930's by the standard works of Allen and Bowley (1935), who used cross-section data to study demand as a function of income, and of Schultz (1938), who used time series data to study demand as a function of income and prices. These estimates of demand functions utilized least squares regression. In recent contributions we note the following points:

(a) The empirical work has continued and led to a reliable mapping of the basic features of the demand structure.† In particular, the results have proved applicable for prediction, as when forecasting the demand for various commodities after a period of shortage or rationing (Wold-Juréen, 1952, Chs. 1.6, 18.2; Fox, 1953, p. 72).

(b) The empirical studies use cross section data and time series jointly (Wold, 1940; Tobin, 1950; Wold-Juréen, 1952; Stone, 1954).

(c) The existence and certain general properties of demand functions having been established on the basis of the Pareto-Slutsky theory of indifference maps, these theoretical results are exploited for testing the empirical demand analysis. The theoretical results include several general relationships between demand elasticities.‡

(d) Schultz's work left unsettled the question whether to use the regression of price X on demand Y or that of demand on price for estimating the demand function. Hence, he calculated both regressions and gave two estimates of elasticity. The argument that the two regressions correspond to two different causal hypotheses, as indicated in (2 a–b), and that the notion of demand function implies hypothesis (2 a), leads to the decision to use the regression of demand on price (Wold-Juréen, 1952, Chs. 1.2–3, 14.2). Since this conclusion is a statistical implication of the basic theoretical model, it may be regarded as a special instance of the interplay between theoretical and empirical inference (referred to under c above). The conclusion extends to partial regression, e.g., the simultaneous estimation of the parameters of a demand function with two constant elasticities, say

$$d = \text{const.} \; \mu^E p^{-e} \qquad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where d is demand, μ income, p price, E income elasticity, and e price elasticity taken with a negative sign in accordance with Marshall's original definition.

* Assuming demand for coffee as a function of price to be $d = A \, . \, p^{-\alpha}$, Benini estimates the parameter α, or in the terminology of Marshall (1890) the price elasticity of coffee, to be $\alpha = 0{\cdot}384$. Based on yearly data for Italy 1880–1893, the estimate is obtained in a quite modern fashion as the regression coefficient of log quantity on log price.

† Several comprehensive studies have appeared. For the treatment of U.S. data, see Shepherd (1941), Tobin (1950), Fox (1953); U.K. data, Stone (1954); Swedish data, Wold (1940), Malmquist (1948), Wold-Juréen (1952). A multicountry study by Juréen is now in the press (*Econometrica*, 1956).

‡ See Wold-Juréen (1952), also for further references.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(iii) *Fertility and reproduction.* The variation of fertility provides a striking instance of a causal complex that has not been mastered operationally. The annual births in Sweden decreased greatly from 1910 to 1934, followed by a 50 per cent. increase up to 1946 and another decrease thereafter (Fig. 1). It is easy to list causal factors behind the 1934–46 rise in births: fuller employ-

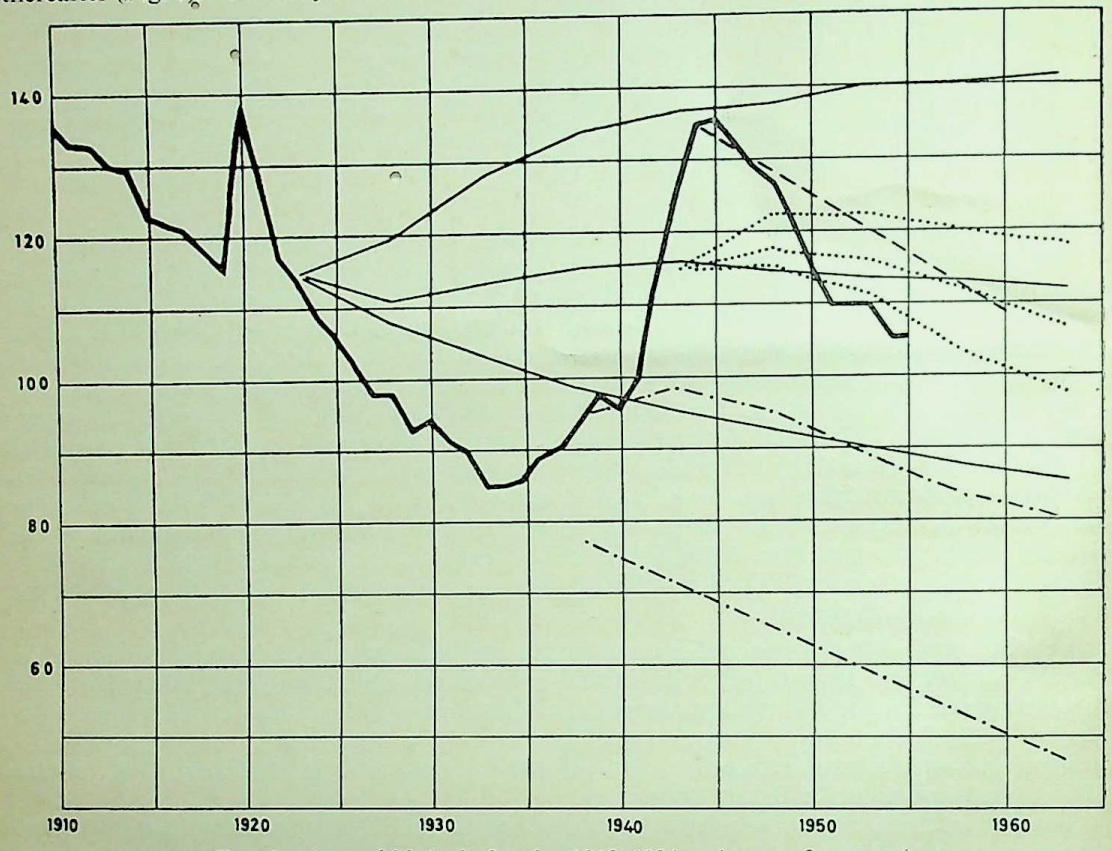


FIG. 1.—Annual births in Sweden 1910–1934 and recent forecasts.*

ment, earlier marriages, increased incomes with more even distribution, decreased abortions, paid holidays at pregnancy, maternity grants, housing subsidies, the shortage of commodities during the war,† etc.; but opinions are divided as to the relative importance of these several factors. The lack of quantitative knowledge about the causal relationships is illustrated by the divergent forecasts of future fertility.

Forecasts have been made on the basis of mechanical extrapolation of death-rates, marriage-rates, fertility-rates combined with high, low and medium estimates for future trends. Since the forecasts rest on no explicit causal analysis of fertility variations it is hardly surprising that they deviate widely from actual developments. Though such forecasts have a certain utility, especially in periods of stable development, they are descriptive rather than explanatory, and typify an intermediate approach between the NW and SW sectors.

2. A REVIEW OF EXPLANATORY APPROACHES, WITH PARTICULAR REGARD TO THE TREATMENT OF EXPERIMENTAL V. OBSERVATIONAL DATA

The material of this section is arranged under three headings, in line with R. A. Fisher's distinction between the consecutive phases of a statistical investigation: specification of hypotheses, estimation problems and hypothesis testing.

* Most of the forecasts are reproduced from official publications; for references, see Wold (1956). The graph has been prepared in collaboration with Professor S. Malmquist.

† It has been pointed out that the expenditures connected with the increase in fertility roughly correspond to the drop in imported cars.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2.1. *The Specification of Hypotheses*

In explanatory approaches we may in general distinguish between two component parts in the model which forms the basis of the statistical treatment of the data, namely, causal assumptions and stochastic assumptions.

A. *Causal assumptions.* These form the main subject-matter content of the explanatory approach. The causal assumptions are usually extracted from prior experience, a frequent case being that they form a special application of a previously established subject-matter theory. In experimental situations the mechanism of the experiment is designed so as to test or demonstrate the causal assumptions.

B. *Stochastic assumptions.* These are part of the subject-matter content of the approach inasmuch as they give an interpretation of the deviations between observed and theoretical or expected values. At the same time they supplement the causal assumptions in providing the rationale of the statistical techniques for estimation and hypothesis testing.

Illustrations:

(a) SE sector. An animal experiment with two or more different treatments, each with a fixed dose and the control group with dose zero, the explanatory hypothesis being that the treatments may have different effects. We consider the statistical treatment of such data by the t-test or by analysis of variance. The causal assumption is then embodied in the parameter which indicates the varying mean for the different treatments; on the null hypothesis the parameter takes the same value for all treatments. The stochastic assumptions specify the frequency distributions of the various treatment effects: the replications are assumed to be mutually independent and the distributions are taken to be normal with the same variance and with means given by the treatment parameters.

(b) SE sector. A dosage-response experiment with three or more steps of dosage. In the treatment of such data by regression analysis, the causal assumption brings out the variation of response with dosage, by way of a regression curve or a parameter β indicating the slope of a linear regression. The stochastic assumptions specify the deviations from the theoretical regression. The deviations are assumed to be mutually independent, and in a randomized experiment they will, as a rule, be uncorrelated to a good approximation with the dosage. In the ordinary t-test deviations are further assumed to be normally distributed with zero mean and constant variance. A typical problem is to estimate the parameter β and to assess the sampling error of the estimate by way of a confidence interval.

(c) SW sector. Demand analysis, say an application of formula (3) to time series data. The causal assumptions are embodied in the relation (3), the interpretation being that an increase of income μ by 1 per cent. or a decrease of price by 1 per cent. will, *ceteris paribus*, cause an increase in demand d by E per cent. or by e per cent. respectively. Expressing the relation in logarithms,

$$\log d = c + E \log \mu - e \log p + z \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where z stands for the deviations between observed and theoretical values, the stochastic assumptions refer to the properties of the residual z. Formally the situation is analogous to a dosage-response experiment. On the assumption that z is uncorrelated with $\log \mu$ and $\log p$, the least square estimates of E and e under general conditions will be consistent. If we add the assumptions that the z-values are independent and normally distributed with constant variance, the sampling error of the estimates can be assessed by the same method as in illustration b.

In the following subsections (i)–(iii) the specification of hypotheses is commented upon from the point of view of experimental *v.* observational data. The first bears upon the relative importance of the causal and stochastic assumptions in various types of research, the second upon the degree of complexity of the causal assumptions, the third upon the systematic co-ordination of the two sets of assumptions.

(i) *The degree to which refined statistical techniques are required.* It happens quite often that the explanatory approach can be carried out by the use of averages, frequency distributions or other elementary devices and that the significance of the results can be judged by common sense without refined techniques. In such cases, of course, the main emphasis lies on the causal assumptions. Experimental situations represent an extreme in the other direction. Generally speaking,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the design of the experiment exploits the technical possibilities to bring the uncontrolled variation under statistical control, by randomization and other devices; this control is brought out in the stochastic assumptions (cf. illustrations a and b above), and these form the basis for the exact assessment of confidence intervals for parameter estimates, significance levels in hypothesis testing, and of other elements of the causal inference.

To illustrate the common sense approach, two striking cases may be cited from the recent Presidential Address of R. A. Fisher to this Society (Fisher, 1953); the geologist Lyell's assessment of the chronological order of fossiliferous strata on the basis of the varying frequency of still surviving species; the biologist Schmidt's observation that the eel differs from other fish in showing no geographic variations in the frequency distribution of the fin-rays, and his ensuing theory that the eels of Europe and the Nile have a common breeding centre in the ocean. In both cases the approach is statistical, but the focus of attention is entirely upon the causal hypothesis. The statistical argument is limited to a comparison between frequencies or frequency distributions which could be performed directly, without the use of any statistical test. Hence it is only natural that the whole credit of the results falls on the subject-matter specialist, whereas the statistical part of the work retires into the background once the causal hypothesis has become established and generally accepted. "The Moor has done his work, the Moor can go", to quote Schiller. The cases of Lyell and Schmidt are rather extreme in this respect, it is true, for, as pointed out by R. A. Fisher, the statistical origin of their theories was almost immediately forgotten.*

Most explanatory approaches based on observational data are ranged between the two extremes commented upon above. Illustration (c) is a typical case. Since the data employed for estimating (4) do not come from a randomized experiment in Fisher's sense, income and price may be correlated with other factors that influence demand, and such correlation will in general impair the interpretation of (3) as specified in the causal assumptions. The resulting deviation between estimated and hypothetical parameters is known as an *error of specification*. The presence of a specification error is not indicated by the standard error of a regression coefficient or by its confidence interval. Such probability margins only account for the sampling error. It is also a familiar fact that no routine methods are available for guarding against a specification error. Instead, *ad hoc* methods of various types come to the foreground. This is treated later on; for the moment we sum up the argument symbolically,

$$\text{sampling error} \longrightarrow \text{specification error} \quad . \quad . \quad . \quad . \quad (5)$$

thus indicating a shift of emphasis when passing from experimental to observational data.

(ii) *The artichoke principle of the experimental method.* A fundamental and well known advantage of the experimental method is that it can proceed step by step in disentangling a complex causal problem, the technical control of causal factors enabling the experimenter to break down the problems into subproblems, each explored by a separate series of experiments. Thanks to this method or principle, the experiments dealing with the partial problems can work with causal assumptions of a relatively simple structure. There are innumerable ways in which this principle is utilized in experimental research. In the systematic co-ordination of experimental design and statistical technique, which is the keynote of R. A. Fisher's work, the principle is well exploited. The pattern of causal assumptions behind the modern statistical techniques is of a simple standardized type covering the basic situations encountered so frequently in experimental statistical work. I would emphasize that as the causal assumptions become simpler the statistician can concentrate more on the stochastic side of the problem; this feature strengthens the argument behind formula (5).

The simplest experimental devices for the step-wise treatment of causal problems are:†

* There are many examples of the same tendency in the literature, and in the scientific workshop any consulting statistician can provide further instances of playing the part of the Moor. This is not meant to discourage consultative work on simple problems. The argument is rather that a statistician with experience in applied work may often render valuable service in the subject field by helping to pose the problems and at the same time making them amenable to statistical treatment. He will then serve both as a statistician and general research consultant, and the conclusion is that his work should be organized in accordance with this double function.

† The various points are more or less explicit in the standard text-books; for an integrated picture of research principles, see Wilson (1952).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(1) Varying one or a few of the controlled causal factors at a time, keeping the other ones constant. (2) Neutralizing the effect of uncontrolled factors by randomization. (3) Arranging for additive and linear effects of the controlled variables. Several devices are available under (3), among these are (a) some appropriate functional transformation of the controlled variables; (b) restricting the variation of a controlled variable to a short interval (so that by Taylor's theorem its influence on the effect variable is approximately additive and linear); (c) restricting the variation of a controlled variable to two values (this restriction is sometimes fulfilled automatically, so-called 0–1 variation), and subjecting the possible non-additivity to a special test of the type known as tests for interaction.

When dealing with observational data the statistician is in a more difficult position. The explanatory approach must then be taken as an entity, and as a consequence the causal assumptions will as a rule be more involved and complicated. As before, it is the absence of randomization that is crucial, for since this device is not available there is no clear-cut distinction between explanatory factors which are explicitly accounted for in the hypothetical model and disturbance factors which are summed up in the residual variation.

The discussion here is closely related to that under (i) and is rather an elaboration of the same general line of argument. As before we state the conclusion as a shift of emphasis in the analysis of experimental *v.* observational data:

$$\text{Elements of causal inference} \longrightarrow \text{Integrity in model building.} \quad . \quad . \quad (6)$$

(iii) *The implications of the stochastic assumptions for statistical technique.* The stochastic assumptions play an essential part in establishing the efficiency of statistical techniques, as regards the accuracy of estimation devices as well as the power of methods for hypothesis testing. Some comment on the common background of estimation and hypothesis testing techniques should now be made.

The standard methods of R. A. Fisher and his followers represent an ideal with regard to the systematic way in which the uncontrolled variation in the experimental replications is exploited for purposes of causal inference. The advantages are manifest both in estimation and hypothesis testing; throughout, the Fisherian methods have a high and even optimal degree of precision and power. This efficiency is the result of a comprehensive specification of the stochastic assumptions. For the simplest cases reference is made to illustrations a–b above. Generally speaking, the optimal accuracy of the modern techniques rests on a fairly complete specification of the distribution of the residual variation. The specification required is not a serious problem in controlled experiments, however, for thanks to the technical design of the replications (the randomization device, etc.) the experimenter can arrange so that the assumptions will be fulfilled in practice, at least approximately. On the other hand, in situations where the explanatory approach is based on observational data there is little or no prior knowledge about the residual variation. Hence its properties cannot usually be specified to the extent required by the modern standard techniques. The consequences of this argument are discussed in detail below.

The modern statistical methods work with stochastic assumptions which are specified with rigour and often in great detail, the basic techniques thus requiring normality, homoscedasticity, etc. Hence it is quite natural that current literature on statistical theory devotes so much space to the task of generalizing and varying the stochastic assumptions behind the standard techniques. The task is so much more important and urgent because of the large array of intermediate cases between observational situations with complete ignorance of the properties of the residual variation and experimental situations where the residual distribution is under sufficient control. The work in this direction has recently been reviewed by J. Tukey (1954) in a stimulating paper. In the present context there are two general arguments of particular interest in Tukey's paper. He points out that the work in question has been almost completely confined to the simple standard situations in regard to the causal hypotheses, with the result that questions referring to systematic *v.* observational errors and other intermediate cases between experimental and observational situations have hardly been touched upon. He also points out that we cannot go far from the simplest types of stochastic assumptions before we are in *terra incognita* with regard to the accuracy and power of the standard techniques and before the exact treatment of the problems becomes troublesome and unwieldy in mathematical and computational respects.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

For the future, it is of course important to continue work on the systematic co-ordination of causal and stochastic assumptions, the goal being an arsenal of accurate statistical techniques from which we may choose for each particular case so that the underlying specification is sufficiently detailed and realistic. Work in this direction has only begun, however, and one must ask whether it can keep pace with the ever-growing number of causal models of varying complexity and their combination with different types of stochastic variation. In any case, the task is enormous if we wish to achieve perfect optimality, since this requires a detailed specification of the distributions entering into the stochastic assumptions. Thus for a long time to come the question, in practice, will be the attainment of an appropriate balance between detailed specification of the stochastic assumptions and accuracy in the resulting technique. Here, a generalization of the assumptions widens the field of potential applications, but this advantage is counteracted by attenuation of the inference. It would seem that large-sample methods of the classic type come near the ideal of such a compromise. Being asymptotically distribution-free they are very flexible, and in ordinary situations they do not diverge far from optimal accuracy. This should be taken as a provisional conclusion, and it will be discussed further in section 2.2–3.

2.2. *Estimation of Parameters and Other Elements of Causal Inference*

With regard to estimation techniques, the discussion in the previous section may be summed up in the following formula for change of emphasis when passing from experimental to observational situations:

$$\text{Methods of optimal efficiency} \longrightarrow \text{Large-sample techniques.} \quad . \quad . \quad (7)$$
$$\leftarrow$$

The small arrow in the opposite direction serves to stress the left-hand side as an ideal, even if in practice we are often far from a perfect co-ordination of causal and stochastic hypotheses and the ensuing optimal efficiency.

For illustration of (7) reference is made to least squares (l.sq.) and maximum likelihood (m.l.) estimation, the principal estimation techniques of classic and modern statistics. The basic facts are well known.* The m.l. method is of more universal scope. Under general conditions of regularity both methods give estimates that are asymptotically unbiased. The l.sq. estimates are consistent, i.e. their standard errors tend to zero with increasing sample size, while the m.l. estimates are even of optimal efficiency, i.e. their standard errors are asymptotically the smallest possible. The higher accuracy of the m.l. method rests on a more detailed specification of the stochastic assumptions; generally speaking the l.sq. method is distribution-free, whereas the m.l. method is not. Note that the l.sq. and m.l. estimates coincide in the simplest situations involving normality linearity, independence, etc.; it is only when dealing with degenerate or irregular cases that the loss of accuracy in the l.sq. estimates becomes considerable. The mathematical and computational simplicity of the l.sq. method should also be mentioned. Hence it is no wonder that the l.sq. method has served and still serves satisfactorily over wide fields of applied statistics: astronomy, geodesy, actuarial science, econometrics, etc.

In the recent development of theoretical statistics, with its emphasis on optimal efficiency, the least squares method has unjustly come into disfavour. The above comments are not intended as criticism of the maximum likelihood method; the point is that both methods have their proper spheres of application, and that the l.sq. method has certain distinct advantages when dealing with observational data. It may be added that each method rests on its own distinctive principle or criterion, and that neither method can be regarded as a surrogate for the other. In accordance herewith, a theory for the two methods can be developed along parallel lines. We note, for example, both are subject to a robustness theorem to the effect that the extremum of the criterion has approximately the same distribution as when the criterion is formed with parameters which extremize the expected value of the criterion.†

* For the m.l. method, see H. Cramér (1945) and the original works of R. A. Fisher (1950). The statements on the l.sq. method refer primarily to the case of linear relations, but by the use of Taylor developments and other devices they extend to more general cases; see Whittaker-Robinson (1924) and Wilson (1952), also for further references.

† For the detailed treatment of a case in point, see Whittle (1954), Ch. 2.9.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In symbols, if $f(x, \theta)$ is the criterion defined by the observations x and one or more parameters θ, let $f^*(x)$ be the extremum of $f(x, \theta)$ in respect to θ, occurring, say, for $\theta^* = \theta^*(x)$, so that

$$f^*(x) = \text{extremum } f(x, \theta) = f(x, \theta^*).$$

Then $f^*(x)$ has approximately the same distribution as $f(x, \bar{\theta})$, where $\bar{\theta}$ is the θ-value extremizing the expected value of $f(x, \theta)$. In fact,

$$f^*(x) = f(x, \theta^*) = f(x, \bar{\theta}) + 0(\theta^* - \bar{\theta})^2 \quad . \quad . \quad . \quad . \quad (8)$$

and under general conditions of regularity the term of order of magnitude $(\theta^* - \bar{\theta})^2$ will be small compared with those fluctuations in $f(x, \theta)$ that are directly caused by the fluctuations in the sample values x.

2.3. *Hypothesis testing*

Tests for frequencies, averages, etc., by large sample approximations to the normal distribution were in current use in the nineteenth century; early developments of testing techniques included in addition Lexis' divergence coefficient and the ensuing definitions of normal, supernormal and subnormal dispersion together with their interpretation in terms of urn schemes under the names of Bernouilli, Lexis and Poisson.* The Fisherian epoch brought rigour in the posing of significance problems, exactness in their treatment, and an intense activity in the construction of tests for various types of hypothesis and (after the impetus given by J. Neyman and E. S. Pearson) in studying the power of the tests to discriminate against counter-hypotheses. The exact assessment of significance levels, power functions, etc., rests on a detailed specification of the stochastic assumptions with regard to the distribution of the residual variation. This is entirely in line with the fact that the modern statistical techniques have been designed in connection with experimental applications, where the residual variation is under sufficient control for the purpose.

With observational data, the modern approach in hypothesis testing is more or less hampered by the scanty knowledge about the properties of the residual variation, and this is particularly true in the treatment of counter-hypotheses. The need for testing the statistical inference is no less than when dealing with experimental data, but with observational data other approaches come to the foreground. The change in the situation has several aspects. J. Tukey has drawn attention to the fact that the modern techniques for hypothesis testing are designed in line with "the tradition of self-contained experiment", the techniques being based solely on the data under investigation (Tukey, 1954, p. 716; Fisher, 1935, esp. p. 69). In the analysis of observational data the treatment of a single sample is more directly tied up with other experience, empirical and theoretical; hence in hypothesis testing various sources of evidence are taken into account by way of checks and cross-checks. I should like to formulate this tendency as follows:

$$\text{Testing on the basis of the sample} \longrightarrow \text{Testing against other knowledge.} \quad . \quad (9)$$
$$\leftarrow$$

Hypothesis testing in the right-hand sense is always present in experimental investigations as well. Hence what is stated in (9) is that testing on the basis of the sample is of limited relevance in observational research, and that the checking of causal inference against other sources of knowledge is more indispensable, overt and explicit with observational than with experimental data. The small arrow serves to stress the modern techniques to the left as an ideal, both with regard to the exactness of the methods and to the possibility of assessing, internally, the reliability of the inference from the data.

The testing or checking of statistical inference on the basis of other knowledge is largely a matter of *ad hoc* devices varying from case to case. We may distinguish between three main sources of knowledge available for hypothesis testing in a broad sense; namely, empirical data, theoretical arguments *a priori*, and predictive tests. These will now be briefly discussed; for illustration we shall mainly refer to various aspects of demand analysis.

(i) *Hypothesis testing by the use of other empirical data.* There is, of course, no end of ways in which the conclusions from the sample under investigation can be compared with and tested against similar or related inference from data referring to other geographic regions, other social

* For a review and development of Lexis' theory, see Charlier (1912).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

strata, earlier periods of time, or to whatever sample might be appropriate for the purpose. The broader and more varied the outlook, the better. The causal inference from different sets of empirical data are like pieces in a puzzle; they should combine to display a consistent total picture, and such an integration amounts to a check or test of the inference that is of essential relevance even if it does not, like the modern testing techniques, lead up to a numerical confidence level (Wold-Juréen, 1952, Ch. 2.5 (v)). We have mentioned one sample as checked against other evidence, but usually the investigation at issue is based on several different samples, and these will then in the first place be utilized in the puzzle. This is so in demand analysis, where a comprehensive study may include the estimation of scores or hundreds of demand elasticities referring to different commodities, different family types, and so on. A feature of general scope is that part of the statistical data in demand analysis are time series, part are cross-section data, and the more or less parallel results from these two sources may be employed for testing purposes in various ways.

The rationale of the right-hand approach in (9) is obvious, but it is equally clear that in practice it is hampered by various frictions, some of which are in the realm of organizational and institutional questions. There is the strong impact of the tradition of the self-contained experiment, a force which is felt widely outside the sphere of experimental research. This impact is strengthened by a marked difference between descriptive and explanatory approaches. The inference drawn from a separate sample will always have a descriptive aspect, and in this respect the inference may be both interesting and important, but it is only when it comes to application outside the sample that the need arises to check and test the inference against other evidence. For example, the regression of food expenditure on earnings in a certain group of households has its value as one of the statistics to describe the standard of living in this particular sample, but if interpreted as a causal relation to explain how food expenditure varies with income it must be carefully checked and tested against other data before it can be applied outside the sample, say for forecasting purposes. In the treatment of observational data it is therefore a temptation to rest content with the descriptive results and to slight or avoid the more difficult and committal explanatory problems. This tendency is accentuated as another aggravating force comes in by the institutional features of scientific research. In their work for examination degrees, in research grants, and in many other situations research workers are required to report definite results, and as a consequence limited projects are at a premium. Time-absorbing projects are out of the question or appear as doubtful investments, and this applies particularly to the large-scale projects that would be required to explore the causal complexes if we wished to obtain an operational grasp on the broad problems in the social field, such as those referred to in the beginning of section 1.4.

(ii) *Hypothesis testing by* a priori *arguments.* Some theoretical model, more or less elaborate, always enters in the specification of hypotheses behind an explanatory approach. Frequently part of the model enters in the specification, whereas other parts may be employed as auxiliary tools in the empirical analysis, and in particular for purposes of hypothesis testing. This is so in demand analysis, where the Pareto-Slutsky theory of consumer demand provides the existence of the demand functions which in the specification of hypotheses are adopted as a starting point for the empirical analysis. The same theory establishes certain properties of demand functions in the form of general relationships between different demand elasticities, and these relations can be utilized as checks or tests upon the empirical estimates of the elasticities involved. Further comments on this point will be made in section 3.3 (i).

The general features of observational research dealt with under (i) and (ii) place the statistician under a double obligation: to incorporate several types of data in the same investigation, and at the same time to co-ordinate the statistical technques with the subject-matter theory of the field of application. In practice these features force the statistician to specialize in one field, or perhaps a few, where he is familiar with the sources of statistical data so as to know their reach and limitation, and where he can acquire a working knowledge of the subject-matter theory. Without such orientation in subject-matter the statistician runs the danger of missing the important problems in applied work, however versatile he may be with the formal aspect of statistical techniques.

(iii) *Predictive tests.* The ultimate test of the causal inference lies in the question: Are the results useful in practice? In particular, can they be used to obtain realistic and efficient forecasts? Confrontation with fresh evidence is, clearly, the real touchstone for scientific inference in general, and in particular for explanatory approaches based on observational statistics.

The problems that challenge research in the SW sector of our map vary immensely in complexity and difficulty, and so does the degree to which they are mastered for purposes of prediction. Mortality and fertility here represent two extremes, as stressed above (1.4, illustrations (i) and (iii)). In the first mortality tables, constructed some 250 years ago, the influence of age upon mor-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tality was brought out with an accuracy that was sufficient for the special type of forecasting involved in the running of life insurance companies. The fertility variations are of vital importance to the welfare of a nation, but the failure of forecasts in this field shows that we are far from a quantitative knowledge of even the main features of the causal complex in an operational sense. Demand analysis is an intermediate case both with regard to difficulty and positive results. The work in the last 30 years or so has settled the basic questions of method for measuring the demand structure; the actual features of the structure are now known in broad outline, and some of the results have stood the test of prediction over relatively short periods. On the foundation thus established, research can now proceed to the study of special and more complex problems concerning consumer behaviour.

The comments in (i)–(iii) might seem anti-theoretical at first sight, but this is actually not so. For example, in a research project for explaining the fertility peak in Fig. 1 the co-ordination between statistical procedures and subject-matter theory would require qualified statistical work. It would be of great practical importance if this could lead to better forecasts, even if the future development far exceeded the tolerance margins set by the sampling errors. However, the difficulties to overcome in such an analysis are radically different from the types of problems specific to the statistical analyses of a self-contained experiment.

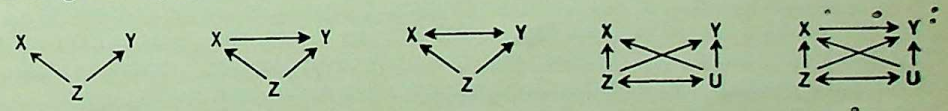
3. The Explanatory Analysis of Observational Data: Some Typical Situations

In order to discuss certain general features in the SW sector of our map we shall now take up four types of approach. These are selected to illustrate the differences in complexity and stage of development in various fields of application, from the first groping attempts to trace a causal relationship that may be taken as a basis for a theoretical model, to the streamlined statistical techniques for routine applications of well-established theoretical models. The intention thus is not to describe various stages in one and the same investigation. The main character of the approach is sometimes empirical, sometimes theoretical, sometimes intermediate, but in the SW sector the statistical method is always closely linked with subject-matter theory, and always seeks to copy the experimental method by *ad hoc* devices, utilizing such features in the actual situation as are analogous in some respect to a genuine experiment.

3.1. *Causal Patterns*

The approaches under this heading are concerned with the existence and direction of causal relationships. A typical problem is to distinguish between cause and effect in a complex of interrelated factors, or to distinguish genuine causal dependence from spurious relations due to common influence from other factors.* The possible hypotheses in the case of two variables are covered by formulae (2 a–d). With the introduction of more variables the possibilities multiply with great speed; among the simplest are:

(10 a–e)



Problems concerning causal patterns are well known from the theory and application of correlation analysis.* If X, Y are intercorrelated, the partial correlation of X, Y with respect to Z yields a test criterion for discriminating (10 a) from (10 b) or (10 c), and by considering the partial correlation with respect to a larger set of variables U, V, . . . we can continue to explore the possibility of spurious correlation. In principle this is a general device. A fundamental approach in this direction is Sewall Wright's method of path coefficients for exploring complex causal patterns; this reference belongs partly to the next subsection since the path coefficients give quantitative information on the relationships.†

* The groundwork is covered by the text-books of Yule and Ezekiel. For a recent review, with special attention to spurious correlation, see H. A. Simon (1954).

† See Wright (1934) and his contribution in the volume edited by Kempthorne *et al.* (1954), also for further references. Tukey's paper in the same volume re-formulates and develops Wright's approach in terms of regression coefficients.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Causal patterns like (2) and (10) can sometimes be explored with advantage by the use of frequency comparisons. An example reported by W. Cochran (1954) deals with the problem of whether there is some hereditary element in the incidence of cancer. Here X and Y are relative frequencies for cancer of the uterus. There are two frequencies X, the first defined for a group A of patients who have the disease, and thus $X = 1$, the second for a group B forming a sample from the total population. Similarly there are two frequencies Y, one defined for a group of specified relatives of A, the other defined for relatives of B. Formulae (2 c and d) indicate the hypotheses of presence and absence of hereditary aggregation of the disease, and formulae (10 c and a) corresponding hypotheses if the data are stratified with respect to some disturbing variable Z; in the present case Z distinguishes between an older and a younger generation of relatives. The difference between the frequencies Y forms a test criterion for discriminating between the two hypotheses, and Cochran shows how this test can be performed by an application of the χ^2-method. It turns out that there is an association of type (2 c), but on stratification this vanishes in the sense of (10 a); the conclusion thus is that the data do not support the hypothesis of hereditary aggregation.

The method employed by Cochran extends to patterns with more stratification. The approach is of further general interest as an instructive design of sampling for purposes of causal inference. Current text-books on the theory and application of sampling deal almost exclusively with sampling techniques that are designed for some descriptive purpose. The optimal allocation of sampling units will, however, be radically different if the data are to serve an explanatory analysis. For example, if we wish to estimate the number of smokers in the total populaton it may be appropriate to use proportional sampling, say after stratification according to income, but if the problem is whether smoking is influenced by income, we should rather select the same number of individuals in each income stratum.* The increase of work in this direction is one of the indications that the SW sector is becoming of more immediate concern to professional statisticians.

If the circumstances are favourable it is possible not only to give evidence about the absence of causal connections, but also to establish the presence and the direction of such relationships. A classic instance is the use of observations on twins for discriminating between hereditary and milieu influences; owing to the close similarity between one-egg twins, they approach the ideal of the replicated experiment under constant conditions (see Dahlberg (1926)). For another case in point, reference is made to a recent study of denture diseases by E. Welander (1955). Caries, tartar incrustations (calculus), soft coatings, pocket formations, mouth hygiene, positional factors in the denture, diet habits, etc. form a causal complex that is hard to disentangle by *a priori* arguments. However, thanks to a coincidence in favourable features, certain aspects of the complex can be studied by statistical methods. The symmetry of the mouth allows us to eliminate the positional factor in the comparisons between different teeth; also, each tooth has 4 or 5 different surfaces, and by taking into account the points of contact of these, their position relative to places of affinity for the various denture diseases, etc., it is possible to specify certain plausible hypotheses on the causal complex. By utilizing and enlarging upon the available Swedish records of the requisite detailed type, Welander found, among other things, that there is inter-individual correlation between caries and calculus, and also between caries and pocket formations; the first correlation vanishes when formed for teeth with symmetric position, but the second does not. This result agrees with the plausible argument that caries and calculus have no local interaction, whereas caries and pocket formations should have.

3.2 *Regression Analysis*

Next we turn to situations where the existence and direction of a causal relation has been established (perhaps tentatively) and we wish to quantify the relation. The problem is thus to estimate the effect variable as a function of the explanatory variables. As regards the causal and statistical interpretation of such a function we note : (a) the relation indicates how the effect will vary as one explanatory variable is allowed to vary while the others remain constant; in general the relation further involves a residual and this accounts for the effect of disturbance factors which

* The problem has been treated from this point of view by Frankel and Stock (1939). Cf. the uniform allocation of sampling units in contingency tables employed in the Kinsey report (1948).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

have not been explicitly introduced as explanatory; (b) the explanation brings back the variations of the effect variable upon variations in the other variables; in general the explanation is partial, the residual variance being a measure of the variation not explained.

The least squares regression of the effect variable upon the explanatory variables yields an estimate of the causal relation, and under general conditions the estimate is consistent. One or two aspects of the situation have been illustrated and commented upon above (sections 1.4 (ii); 2.1, illustration c; and 2.2). Regression methods being of fundamental importance in the SW sector, we shall now discuss the approach in some detail.

(i) *The least squares principle.* As is clear from (a) and (b) above this principle constitutes an obvious rationale of least squares regression when the problem is to estimate a causal relationship. In fact, when applied to the given sample, no other estimate of the relation will give such small residual variance. For a linear regression, as is well known, this minimum property of the residual is equivalent to its property of having (in the given sample) zero correlation with each of the explanatory variables (e.g. Wold-Juréen, 1952; theorem 12.1.1–2). Referring to (a) above, this feature allows us to formulate the rationale of the regression in another way, namely, as being unbiased on the assumption that the joint influence of the unspecified factors is uncorrelated with the explanatory variables.

(ii) *Specification error.* The estimation of a causal relation by regression analysis on the basis of observational data is formally the same procedure as in the case of experimental data, but as stressed in formula (5) there is an increased risk of specification errors, since the uncontrolled factors cannot be neutralized by randomization. Let

$$y = f(x^{(1)}, \ \ldots \ , x^{(h)}) + z \text{ or in the linear case } y = b_1x^{(1)} + \ \ldots \ + b_hx^{(h)} + z \quad \text{(11a–b)}$$

be the hypothetical relation, and suppose that we are primarily interested in how y is influenced by $x^{(1)}$ and $x^{(2)}$. How many and which further explanatory variables $x^{(3)}$, $x^{(4)}$, ... should be introduced is essentially a subject-matter question. This is a most important question, for the neglect of a relevant factor which is correlated with $x^{(1)}$ or $x^{(2)}$ will result in a specification error which will not tend to zero with increased size of the sample; on the contrary, it will remain of the same order of magnitude as the regression coefficient to be estimated. The dilemma is that the factors of potential influence are often numerous, whereas in practice the regression analysis must be confined to a limited number of explanatory variables. The traditional compromise is to include as explanatory those variables which, according to experience and *a priori* theory, are believed to be the main causal factors, the resulting regression being tested for sampling error by a significance formula of ordinary type, and for specification error by one or more *ad hoc* devices, (in line with the comment in section 2.3, (i)–(iii)).

A formal analysis of what can happen when another explanatory variable $x^{(h+1)}$ is introduced in the linear regression (11 b) elucidates the dilemma at issue. We note:*

A. In principle, $x^{(h+1)}$ may upset the regression picture entirely, for there exist variables $x^{(h+1)}$ with unit variance such that a given coefficient, say b_1, will change to an arbitrarily prescribed value b between $-\infty$ and $+\infty$.

B. All coefficients b_i will remain the same if $x^{(h+1)}$ is uncorrelated with every $x^{(1)}, \ \ldots \ , x^{(h)}$. More generally, b_1 will remain the same if $x^{(h+1)}$ is uncorrelated with the residual in the regression of $x^{(1)}$ on $x^{(2)}, \ \ldots \ . x^{(h)}$.

C. Suppose that y allows the representation

$$y = \beta_1x^{(1)} + \ \ldots \ + \beta_hx^{(h)} + z^* \quad . \quad . \quad . \quad . \quad (12)$$

with a disturbance term z^* which (a) has small standard deviation, say $\sigma(z^*) < \theta_1$, and (b) has small correlation coefficients with $x^{(1)}, \ \ldots \ , x^{(h)}$, say $r(z^*, x^{(i)}) < \theta_2 \, . \, \sigma(z^*) \, \sigma(x^{(i)})$. Then the regression coefficients in (11 b) approximate the true values β_i in (12) within margins that are small of order $\theta_1\theta_2$,

$$| \, b_i - \beta_i \, | < c \, . \, \theta_1\theta_2 \ (i = 1, \ \ldots \ , h) \quad . \quad . \quad . \quad . \quad (13)$$

where c does not depend upon z^*.

* For proofs and further comment on the theorems, see Wold-Juréen (1952), Theorems 12.1.3 and 12.3.2–4. See also Ezekiel (1930), p. 347.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In C we have a proximity or robustness theorem for least squares regression which lends partial support to the traditional compromise. Stated in words, if the residual is small and the additional variable has no more than a slight correlation with the other explanatory variables, the two conditions strengthen each other so as to keep the specification error within limits of a smaller order.

(iii) *Sampling error.* For regression coefficients, just as for most statistical estimates, the sampling error is of the order $0(1/\sqrt{n})$, if n is the size of the sample. Thus in large samples it tends to be of a smaller order of magnitude than the specification error. There is always the danger of specification error in a regression relation based on observational data, so when assessing its sampling error in the form of a standard error or a confidence interval, the results will make only a partial test of the validity of the regression. This is by no means a recommendation to ignore the sampling errors in applied work in the SW sector of our map; what I wish to stress is that they must be supplemented by validity tests of the type discussed in 2.3, (i)–(iii).

The many regression variants and the corresponding tests of significance provide a typical illustration of a general feature commented upon in 2.3, namely that the exact test techniques require a detailed specification of the stochastic assumptions. The situation was briefly discussed in connection with formula (4). In the standard tests it is assumed that the residual variation is homogeneous in the sense of forming sample values from one and the same distribution and that these sample values are independent (see Cramér, 1945, Chs. 29.8, 29.12, 37.2–3). This type of assumption is well suited for experimental situations, independence and homogeneity of the replications being essential aims in the experimental design. With observational data, on the other hand, these assumptions are often impaired by tendencies to patchiness and cluster, a characteristic feature of social and economic material. One source of such interrelation lies in the slow diffusion of behaviour patterns, which gives rise to similarities between sampling units that are neighbouring in space or time, to group effects within different social strata, etc. Here we have entered the danger zone between sampling error and specification error. At the same time we note that, provided the interrelations follow some regular pattern, they need not in themselves imply a specification error or a change in the order of magnitude of the sampling error. Reference may here be made to the regression based on time-series data of the stationary type. The interrelation then takes the form of autocorrelation in the residual series, and this feature can be accounted for in assessing the sampling error. Thus in a linear regression (4) with one explanatory variable the standard error of the regression coefficient has to be corrected by the multiplicative factor

$$\sqrt{(1 + 2r_1\rho_1 + 2r_2\rho_2 + \ldots)}\ , \qquad . \quad . \quad . \quad . \quad (14)$$

where $r_1, r_2, \ldots$ and $\rho_1, \rho_2, \ldots$

are the autocorrelation coefficients of the explanatory variable and the residual, respectively.* We see that the large sample standard error remains of the order $0(1/\sqrt{n})$.

(iv) *Observational error.* A complication arises if the explanatory variables are subject to inaccuracy in measurement. Corrections for such error have been worked out on the assumptions that the errors are independent of the error-free variables and that the error variances are known *a priori* (e.g. Wold, 1940, §16). In general the correction is of the same order as the regression coefficient itself; thus, like the specification error, it is of a larger order of magnitude than the sampling error. The disadvantage in such correction methods is that in practice we have little or no information about the observational errors, neither of their presence nor of their distributional properties. Hence if it is felt that the data contain observational errors that are not negligible (negligible relative to the influences upon the effect variable that are due to unspecified causal factors), it is dangerous to employ the data for more than a tentative orientation.

A problem sometimes dealt with in statistical literature refers to the situation in which the relationship to be estimated is exact in the mathematical sense, so that the residual deviation is entirely due to the presence of observational errors.† This type of situation is well known from astronomy, geodesy, and related sciences. For example, when determining a comet's orbit on the celestial globe the latitude-longitude observations are subject to error, and the orbit is treated

* For the proof of this and more general results, see Wold (1950) and, also for applied work, Wold-Juréen (1952), Chs. 13.4, 15.2 and 17.2–7. Further results in this direction are due to Grenander (1954), who shows, *i.a.*, that there are cases where a least squares estimate has optimal efficiency, asymptotically, in spite of the autocorrelation.

† For a recent review, see Durbin (1955). Cf. also Wold-Juréen (1952), Ch. 2.4.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

as an exact relation between the error-free observations. For such applications an adjustment technique is available, based on the least-squares principle. Within their limited field of application these techniques are indispensable; on the other hand it can be said that the statistical aspect of these matters is of secondary importance, for in practice the main concern in regard to observational errors is to reduce them by improved methods of measurement, and especially to reduce systematic errors.

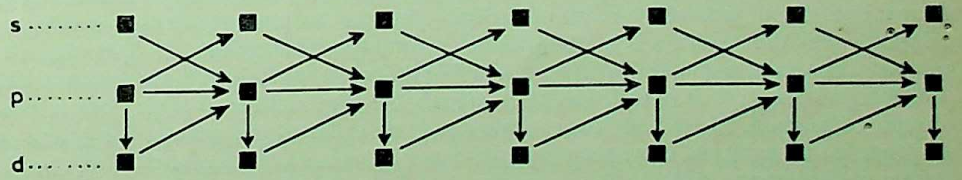
(v) *Simultaneous regressions.* A regression relation which has been obtained as an estimate for a hypothetical causal relation may in principle be employed just as if it were the theoretical relation itself, for any purpose within the limits posed by general logic and special subject-matter theory. Thus it can be combined with any other knowledge, theoretical or empirical, and in particular it can be put together with other regressions to make models in the form of simultaneous relationships. For pioneering applications of such models reference is made to the work of J. Tinbergen (1939) on the statistical testing of business cycle theories. Because his approach is instructive from the point of view of ends and means in causal inference it will be considered in some detail. The systems of Tinbergen are built so as to allow a clear-cut causal interpretation. Each relation refers to a specified *economic unit* and serves to explain its actions or behaviour with regard to a specified variable; for example, a demand relation shows how the group of consumers change or adjust their demand according to changes in price, income and other explanatory variables. The variables thus explained by the system are called *endogenous*; other variables appearing in the system are called *exogenous.* The economic units and their behaviour patterns are *autonomous* in the sense that any unit may modify or alter its behaviour pattern without affecting the behaviour patterns of the other economic units. Finally, the relations are specified as a dynamic system in which the causal relations link together so as to form a *causal chain*; that is, on the basis of the past history of the system, its relations enable us to predict the future path of the effect variables recursively, one by one, and period by period. Systems of this type are known as recursive systems or *causal chain systems.*

The following simple system is a theoretical demand-supply model of causal chain type,*

$$d_t = D(p_t);\ s_t = S(p_{t-1});\ p_t = p_{t-1} + \gamma(d_{t-1} - s_{t-1}) \text{ with } \gamma > 0 \quad . \quad . \quad (15a\text{–}c)$$

The economic units are the consumers for the demand relation, the producers for the supply relation, and the merchants, who in a free market serve the function of bringing demand into contact with supply and of regulating the price mechanism, for the price formation relation. In (16) the causal chain is illustrated by an *arrow scheme* of the type introduced by Tinbergen; it indicates, for example, that supply at period t is assumed to be a function of price at period $t-1$. The causal chain as pictured by an arrow scheme forms a dynamic generalization of the causal patterns in (2) and (10). Two stipulations that general logic imposes on the arrow scheme are (a) it must not be ambiguous or self-contradictory in the sense that there are two or more

(16)



different causal relations for one and the same variable, and (b) a variable that occurs as causal with a certain lag must not appear later on in the chain as effect variable with the same lag. If such a causal circle were admissible, some endogenous variable via one or more intermediate causal links would form part of its own cause, thus defeating the general purpose of a causal model, namely to express a given variable in terms of others. Generally speaking, the model will be a causal chain system if the relations can be arranged in such manner that all arrows in the scheme point either from left to right or vertically in one direction, say downwards.

The analogy between economic models and physical feed-back systems has recently been developed by A. Tustin (1953). A crucial point is that the physical systems work with feed-backs

* See Wold-Juréen (1952), Chapters 1.4 and 3.2, also for comments and references.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that are so rapid that they can be treated as instantaneous, whereas in causal chain systems the feed-back reaction (the lagged dependence of a variable upon itself) cannot be expressed for a shorter time than the unit of observation. Over large sectors of economics the reactions take place slowly, and then the causal chain systems will be appropriate; in other cases the reactions are rapid and may as a consequence require instantaneous feed-back models.

As regards the statistical treatment of causal chain systems we have already stated that they can be estimated by forming the least squares regression for each relation in the system. In fact, the bias in this regression is a matter of intercorrelation between the theoretical residuals and the explanatory variables, and the situation is the same whether the relation is single or belongs to a system. Naturally, since a system contains several relations, the risk of specification error multiplies, but a specification error in one relation need not affect the others. For example, the price formation in a free market being a complicated mechanism, relation (15 c) may at best be taken as a first crude approximation, but a specification error here is of no consequence in the demand and supply relations of the system.

In estimating causal relations by regression analysis the logical necessity to avoid causal circles makes a clear-cut departure from regression analysis as used for descriptive purposes. Thus if we form a system (17 a) of least square regressions,

$$\left.\begin{aligned} x^{(1)} &= b_{12}x^{(2)} + b_{13}x^{(3)} + z^{(1)} \\ x^{(2)} &= b_{21}x^{(1)} + b_{23}x^{(3)} + z^{(2)} \\ x^{(3)} &= b_{31}x^{(1)} + b_{32}x^{(2)} + z^{(3)} \end{aligned}\right\} \qquad \left.\begin{aligned} x^{(1)} &= u^{(1)} \\ x^{(2)} &= c_{21}x^{(1)} + u^{(2)} \\ x^{(3)} &= c_{31}x^{(1)} + c_{32}x^{(2)} + u^{(3)} \end{aligned}\right\} \quad . \quad (17\text{ a–b})$$

these so-called elementary regressions may always serve to describe the sample on which they are based. For example, the first relation shows how $x^{(1)}$ varies, on the average within the sample, if $x^{(2)}$ and $x^{(3)}$ are allowed to vary, and in the variance of the residual $z^{(1)}$ we have a measure for the accuracy of the description. On the other hand, a regression system of type (17 b) is in the form of a causal model; in fact, for causal circles to be absent, irrespective of the number of relations and variables involved, it is clearly a necessary and sufficient condition that the relations can be arranged so as to make the coefficient matrix triangular. One way to characterize the situation is that in systems of type (17 b) the residual variances are brought down to a minimum of what a causal explanation can achieve with the variables considered. If a system (17 a) were adopted as a causal model, the residuals would be smaller, but th reduction would be due to a causal circle, the variations in a variable being partly explained in terms of themselves. In this quantitative sense, systems of type (17 a) would, as causal models, "explain too much".

Some fifteen years ago the method of least squares regression was vigorously attacked in connection with a study of theoretical models in the form of simultaneous relations, and it was shown that for a general category of models—later to be known as systems of interdependent relations—the l.sq. method is not applicable without bias (Haavelmo, 1943).* The situation was reviewed in a joint paper by R. Bentzel and Wold (1946); we found that for causal chain models with normal and mutually independent residuals the l.sq. method is equivalent to the maximum likelihood method. In a later paper I have pointed out that the general conditions for the l.sq. estimates to be unbiased remain the same and can be satisfied if the estimated relation belongs to a causal chain system (Wold 1951). In regard to the rationale of interdependent systems I should now like to comment from the point of view of description *v.* explanation. A special case, a demand-supply model borrowed from Girshick and Haavelmo (1947), will be cited. With a slight change in arrangement for easier comparison with (15) the model is

$$\left.\begin{aligned} d_t &= \alpha p_t + \beta + u_t \\ s_t &= h_1 p_t + k_1 + v_t \\ d_t &= s_t = x_t \end{aligned}\right\} \qquad \left.\begin{aligned} x_t - \alpha p_t &= \beta + u_t \\ x_t - h_1 p_t &= k_1 + v_t \end{aligned}\right\} \quad . \quad . \quad . \quad (18\text{ a–b})$$

* For later contributions see Hood-Koopmans, eds. (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Both the causal chain model (15) and the interdependent system (18a) are attempts towards the dynamization of Cournot's static model for demand-supply equilibrium. Whereas (15) aims at a complete dynamization, a causal relation being specified for each of the variables involved, the dynamization in (18) is partial, the demand and supply relations being specified while price is dealt with as an equilibrating variable. Hence model (18) stands in close relation to the approach known as *comparative statics*. This is an attractive aspect, inasmuch as the problems of comparative statics are of great practical importance, and the model focuses upon these without entering into the price formation process with its notoriously complicated mechanism. Another interesting feature of interdependent models is that they involve identities which are introduced by way of assumptions, definitional relations, or otherwise, such as $d_t = s_t = x_t$ in (18) or, to give another example, total income = consumption + investment. The identities form a constraint for the effect variables, and this implies that the relations are not autonomous in the sense of causal chain systems (hence the term interdependent systems). Another possible interpretation of the identities is that they represent instantaneous feed-backs in the model. The properties mentioned have no doubt combined to arouse interest in the interdependent systems, particularly with regard to the use of total categories in economic models.

Against these attractive features there is, however, some obscurity in the approach of interdependent relations, especially with regard to the operational significance of the model. Once its parameters have been estimated, what applications can be made of the model? Can a relation in the system be used for estimating one variable in terms of the other ones? According to my understanding this approach may lead into the pitfall of "explaining too much," although the situation is more sophisticated than in the elementary regressions (17 a). The regressions (17 a) refer to the total distribution of the variables involved, whereas in interdependent relations of type (18 b) the left-hand expressions usually refer to the conditional distribution defined by regarding the predetermined variables appearing in the right-hand members as fixed. Further, the left-hand expressions are not defined as elementary regressions of the conditional distribution, but as having residuals with certain specified correlation properties. The number of parameters to the left in (18 b) is, however, the same as in a set of elementary regressions, and this feature gives rise to the risk that the relations "explain too much" if interpreted in a causal sense. In the applications of interdependent systems this risk has actually materialized, inasmuch as some of the reported models have given residuals with smaller variance than is compatible with a causal interpretation of the relations (Wold, 1955). When this happens it implies that the parameters of the interdependent system are, on the whole, larger in absolute size than in a causal model for the same phenomena.*

(vi) *The subject-matter background.* As a tool for causal inference from observational data regression analysis is by no means a routine affair, and above all it is important for the regression to have a sound theoretical background. In experimental situations the design of the experiment indicates how to specify the regression, and in particular it is clear which variables are effects and which are causes. With observational data the specification of hypotheses is in essence a subject-matter question. This is, however, only one aspect of the co-ordination between statistical analysis and subject-matter theory; we shall now proceed to a more general discussion.

3.3. *Model Building on the Basis of Elementary Assumptions*

Causal inference is an interplay between empirical and theoretical analysis, and in this mixed procedure the lead is sometimes taken by the empirical, and sometimes by the theoretical approach. Earlier in this section we have been concerned with only one aspect or function of the theoretical element in the interplay; namely, to pose an articulate question to the data under investigation or, otherwise expressed, to specify the causal hypothesis behind the statistical analysis. Following up the comment in 2.3 (ii) we shall now turn to situations where a theoretical model is available which is only partially exploited in the specification of hyptheses, leaving other parts as ancillary information that can be utilized for supporting and testing the statistical inference. There are several ways in which the statistical analysis can profit in such situations, notably:

* In agreement with this argument, some applications of interdependent systems have resulted in demand elasticities which are larger than those obtained by least squares methods; see Bergstrom (1955), p. 265, and Hildreth-Jarrett (1955), p. 115.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(a) Several types of data can be treated on the basis of the same theoretical model.

(b) The empirical results can be tested for agreement with theoretical relationships, asymptotic properties, etc., that are indicated by the ancillary information.

(c) The rationale of the statistical procedure can be examined by subjecting the empirical results to tests of type (b).

If the statistical hypotheses in this way form part of a more embracing theoretical model, a general feature in the situation is that the model is based on more fundamental and elementary hypotheses. The more firm and rich in substance the theoretical model, the more favourable is the situation for judging the validity of the statistical results at issue. The best examples of well-established models belong to the experimental sciences, as in genetics or statistical mechanics. Such reliable and comprehensive models are an ideal also in the social sciences, but unfortunately difficult if not impossible to attain. Here the models are far less exact, and are usually fragmentary, covering only a limited sector of the field of application. To illustrate the points (a)–(c) above, two models will be cited from the social sciences.

(i) *The Pareto-Slutsky theory of consumer demand.* The fundamental assumption is that the consumer has a consistent pattern of preferences with regard to the commodities under consideration. Under general conditions of regularity this simple assumption suffices to establish the existence of the consumer's demand functions; these indicate how his demand for the various commodities is causally influenced by the commodity prices and by his disposable income.* This basic result gives a common theoretical model for the statistical treatment of cross section data (demand as a function of income) and time-series data (demand as a function of prices and income); we have here an application of type (a). To mention one advantage of this, an elasticity estimated from one category of data may, in principle, be used as a known entity when treating another category of data (so-called conditional regression) (Wold, 1940, §10; Wold-Juréen, 1952, Ch. 2.6).† Furthermore the Pareto-Slutsky theory leads to a number of relationships between demand functions, demand elasticities, etc., propositions that allow applications of type (b). Hereunder we note Slutsky-Schulz' relation involving all elasticities for one commodity; Hicks-Allen's relation involving one type of elasticity for all commodities; Hotelling-Juréen's relation involving a pair of cross elasticities; the theorem of Törnqvist, confirming and extending an empirical conjecture by the author, that for necessities the income elasticity is smaller than the price elasticity, and conversely for luxuries. Finally, we note an application of type (c). The income elasticity for an aggregate commodity should theoretically be a weighted average of the elasticities for the separate commodities; the same relation should hold for the estimated elasticities, and so it does for estimates obtained by least squares regression but not for those obtained by orthogonal or diagonal regression; thus, the l.sq. regression has passed the test without bias, but not the two other estimation methods.

(ii) *Migration models.* In the study of internal migration it is a natural approach to represent the frequency of migration between two places as a function of their distance, say a power of their distance; in symbols

$$f = M(d) \qquad f = A \,.\, d^{-\alpha} \quad . \quad . \quad . \quad . \quad . \quad (19\ \text{a–b})$$

For data on the migration to a district in Northern Sweden, formula (19 b) has given a fairly good fit.‡ The parameter α turned out to be remarkably stable, 4 different communities during 8 different decades giving 32 values ranging between $\alpha = 1{\cdot}72$ and $\alpha = 2{\cdot}87$. These results are pleasing, but, on the other hand, the model (19) is rather formal, an approach intermediate between formal description and causal explanation. A model penetrating deeper into the causal mechanism of the migration movements has recently been presented by R. Porter.‡ The driving force behind the migration is here assumed to be economic, unemployed people seeking the nearest vacant job and employers seeking the nearest people for their vacant jobs. Briefly stated, the migration is thus interpreted as a *matching process.* The mathematical analysis of the process leads to a representation of the distance function in terms of regional integrals over the population density.

* For an exposition of the Pareto-Slutsky theory covering the statements here and in the following, see Wold-Juréen (1952), Chs. 4.5, 6.1–5, 14.2 and 17.2–3.

† For an extension of the approach, see Durbin (1953).

‡ The applications of (19b) are quoted by Mr. Lövgren. I am indebted to Mr. Lövgren and Dr. Porter for placing their as yet unpublished work on migration at my disposal.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The resulting formula gives a good fit to the same data, especially in view of the radical reduction in the number of parameters employed; these are 2, against 64 when applying formula (19 b). A beautiful feature of Porter's model is that it gives an interpretation of the parameter α in (19 b). If we imagine a homogeneous country in the shape of a long line (as Chile) the asymptotic value for the parameter would be $\alpha = 2$, whereas a large circular country (as pre-war Germany) would give $\alpha = 4$. The empirical α-values obtained for Sweden have $\alpha = 2 \cdot 3$ for mean value, and this corresponds quite well to the oblong shape of the country. For another thing, the basic assumptions of the model can be varied so as to explore other hypotheses about the causal mechanism. We may, for example, assume that employees seek the nearest job, while employers seek over the whole country when filling vacancies; this would lead to another expression for the distance function. Moreover, on the introduction of auxiliary hypotheses Porter's migration model can be used for the analysis of several types of data, e.g. the percentage of people who live in the same district where they were born.

3.4. *Joint Treatment of Causal and Stochastic Assumptions by Exact Methods*

The models dealt with in section 3.3 are constructed with exclusive regard to the causal aspects, the deviations from reality being left aside simply as an unexplained residual. We shall now turn to models which have a higher aspiration level, that is, models which are complete in the sense that they specify both the causal and the stochastic assumptions. The two types of models will be referred to here as I and II. Examples of type II are the birth and death process, various epidemic models, models for the learning process, to cite only a few instances from the current literature.*

These two lines of approach represent two important fronts in contemporary research. Speaking broadly, type I emphasizes application, synthesizing several sources of data into an integrated approach, and the aim is to achieve operational results, even though the approximations are crude. In type II the emphasis is on pure research, the approach is more analytic, more exact, and the statistical tests follow the principles of the self-contained experiment. In practice the degree of difficulty determines the aspiration level, and this varies widely, depending upon the field of application. For example, in the complex problems of the social sciences operational results are frequently not available, and therefore a model of type I that actually works will represent a significant advance, even if the residuals are large. In other fields, and here examples could be cited from industrial, biological and medical applications, work on models of type II has already given results of practical use. On the other hand there is a wide gap between the problems solved in the birth and death process and the problems encountered in population forecasts (cf. Fig. 1), and there are also other obvious instances of the same type of shortcoming.

It would seem that these two research fronts display differences in institutional background. The work on models of type I has its main roots in the NW sector of our map, and accordingly is carried on in keeping with the attitudes and traditions of the NW sector. Research of type II is largely inspired by the exact methods of the SE sector, the purpose being to perform a systematic co-ordination between the causal and stochastic elements of the model. At the present time progress in the amorphous SW sector seems to be hampered by the lack of contact between the two lines of approach. From the foregoing it may be seen how each of the approaches has something to learn from the other. Research carried out with the standards of the SE sector will benefit from the broader outlook towards problems which is customary in the NW sector, and in particular it is desirable to cover several types of data by the same model. Approaches originating in the NW sector will profit from posing the problems more vigorously in causal terms, and from more determined efforts to subject the inference to predictional and other decisive tests. The plea for such a mutual exchange has been an underlying incentive for this paper, and my purpose will be fulfilled if the suggestions here presented can contribute to a clearer understanding of the ends and means of present-day statistics.

References

ALLEN, R. G. D. & BOWLEY, A. L. (1935), *Family Expenditure*. London: Staples.
BARTLETT, M. S. (1955), *An Introduction to Stochastic Processes with special reference to Methods and Applications*. Cambridge: Univ. Press.
BENINI, R. (1907), *Giorn. Econ.*, 35, 1053.

* For a recent treatment of this subject, see Bartlett (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

BENTZEL, R. & WOLD, H. (1946), *Skand. AktuarTidskr.*, **29**, 95.
BERGSTROM, A. R. (1955), *Econometrica*, **23**, 258.
BRAITHWAITE, R. B. (1953), *Scientific Explanation*. Cambridge: Univ. Press.
CHARLIER, C. (1912), *Ark. Mat. Astr. Fys.*, **7**, no 17.
COCHRAN, W. (1954), *Biometrics*, **10**, 417.
CRAMÉR, H. (1945), *Mathematical Methods of Statistics*. Stockholm: Geber. Princeton: Univ. Press, 1946.
DAHLBERG, G. (1926), *Twin births and Twins from a hereditary Point of View*. Stockholm : Tiden.
DIETL, J. (1849), *Der Aderlass in der Lungenentzündung*. Vienna: Kaulfuss & Prandel.
DURBIN, J. (1953), *J. Amer. Statist. Ass.*, **48**, 799.
—— (1955), *Rev. Inst. Int. Statist.*, **22** (1954), 23.
EZEKIEL, M. (1930), *Methods of Correlation Analysis*. 2nd ed. 1941. New York: Wiley.
FISHER, R. A. (1935), *The Design of Experiments*. 7th ed. 1949. Edinburgh: Oliver & Boyd.
—— (1950), *Contributions to Mathematical Statistics*. New York: Wiley.
—— (1953), *J. R. Statist. Soc.*, A, **116**, 1.
FOX, K. A. (1953), *The Analysis of Demand for Farm Products*. *U.S. Dept. of Agriculture*, Technical Bull. no. 1081.
FRANKEL, L. R. & STOCK, J. S. (1939), *Ann. Math. Statist.*, **10**, 288.
GIRSHICK, M. A. & HAAVELMO, T. (1947), *Econometrica*, **15**, 79.
GRENANDER, U. (1954), *Ann. Math. Statist.*, **25**, 252.
HAAVELMO, T. (1943), *Econometrica*, **11**, 1.
HALD, A. (1952), *Statistical Theory with Engineering Applications*. New York: Wiley.
HILDRETH, C. &. JARRETT, F. G. (1955), *A Statistical Study of Livestock Production and Marketing*. New York: Wiley.
HOOD, WM. C. & KOOPMANS, TJ. C., Eds. (1953), *Studies in Econometric Method*. New York: Wiley.
KEMPTHORNE, O., *et al.*, Eds. (1954), *Statistics and Mathematics in Biology*. Ames, Iowa: Iowa State Coll. Press.
KINSEY, A. C., POMEROY, W. B. & MARTIN, C. E. (1948), *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
LAZARSFELD, P. F., Ed. (1954), *Mathematical Thinking in the Social Sciences*. Glencoe, Ill.: The Free Press.
LÖVGREN, E., *Migration and the Mobility of Labour*. *Geografiska Annaler*. To appear.
MALMQUIST, S. (1948), *A Statistical Analysis of the Demand for Liquor in Sweden. A Study of the Demand for a Rationed Commodity*. (Doctorial thesis.) Uppsala.
MARSHALL, A. (1890), *Principles of Economics*. London: Macmillan.
MOORE, H. L. (1919), *Pol. Sci. Quart.*, **34**, 546.
—— (1925), *Quart. J. Econ.*, **39**, 357.
PORTER, R., Approach to Migration through its Mechanism. *Geografiska Annaler*. To appear.
SHEPHERD, G. S. (1941), *Agricultural Price Analysis*. 3rd ed. 1950. Ames, Iowa: Iowa State Coll. Press.
SHEWHART, W. A. (1931), *Economic Control of Quality of Manufactured Product*. New York: van Nostrand.
SCHULTZ, H. (1938), *The Theory and Measurement of Demand*. Chicago: Univ. Press.
SIMON, H. A. (1954), *J. Amer. Statist. Ass.*, **49**, 467.
STONE, R. (1954), *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom* 1920–1938, 1. Cambridge : Univ. Press.
TINBERGEN, J. (1939), *Statistical Testing of Business Cycle Theories, II. Business Cycles in the United States of America* 1919–32. Geneva: League of Nations.
TOBIN, J. (1950), *J. R. Statist. Soc.* A, **113**, 113.
TUKEY, J. (1954), *J. Amer. Statist. Ass.*, **49**, 706.
TUSTIN, A. (1953), *The Mechanism of Economic Systems*. Cambridge, Mass.: Harvard Univ. Press.
UNESCO (1954), "Mathematics and the Social Sciences". *Int. Soc. Sci. Bull.*, 6, no. 4.
WELANDER, E. (1955), *The Occurrence of Dental Caries in the Permanent Dentition*. Stockholm: Odontological Inst.; and Uppsala: Univ. Inst. of Statistics.
WHITTAKER, E. T. & ROBINSON, G. (1924), *The Calculus of Observations*. 4th ed. 1944. London: Blackie.
WHITTLE, P. (1954), Appendix 2 in H. Wold, *A Study in the Analysis of Stationary Time Series*. 2nd ed. Stockholm: Geber.
WILSON, E. B. (1952), *An Introduction to Scientific Research*. New York: McGraw-Hill.
WOLD, H. (1940), *The Demand for Agricultural Products and Its Sensitivity to Price and Income Changes* (Swedish). Stockholm (Statens Offentliga Utredningar, 1940 : 16).
—— (1950), *Bull. Inst. Int. Statist.*, **32**, I, 145 and II, 277.
—— (1951), *Sankhyā*, **11**, 205.
—— (1954), *Econometrica*, **22**, 162.
——(1955), *Cahiers Séminaire d'Econométrie de R. Roy.*, **3**, 81.
——(1956), *Nordisk Statistisk Skriftserie*, No. 3 (1955) 95.
WOLD, H., in association with L. Juréen (1952), *Demand Analysis: A Study in Econometrics*. Stockholm: Geber; and New York: Wiley, 1953.
WRIGHT, S. (1934), *Ann. Math. Statist.*, **5**, 161.
YULE, G. U. (1911), *An Introduction to the Theory of Statistics*. 8th ed. 1927. London: Griffin.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Discussion on Professor Wold's Paper

Professor Richard Stone: It is a very great pleasure to have with us this evening Herman Wold, who has contributed so much in recent years to the development of statistical and economic analysis, and to hear his views on the broad subject of causal inference from observational data. With the development of econometrics, this subject has taken on a renewed importance for economists and parallel instances could doubtless be cited from other fields of application.

I am sure that Professor Wold is right to emphasize the importance of substantive theory in all attempts at causal inference. My impression is that this fundamental point deserves emphasis in connection with the analysis of experimental as well as of observational data. The importance of a sophisticated and well-integrated theory seems to me to be associated more with the stage reached in the study of a subject than with the question of whether empirical information is obtained by controlled experiment or by watching the course of events. In the early stages of inquiry it will usually be sufficient to work with rather simple hypotheses in order to get an outline of what is going on, but as experience grows it becomes more and more important to have an integrated theory which will help to resolve the apparent contradictions of earlier, partial investigations. It is a feature of economics with its tendency, historically, to divide into mainly theoretical and mainly descriptive studies that it has islands of sophisticated theory which, though of great use to econometricians, go further than is really necessary in certain directions and not far enough in others.

A fundamental distinction in the paper before us is between models which involve mutual dependence within a period and those in which the variables can be related by a system of causal chains. Professor Wold seems to suggest that the actual world should always be capable of representation by systems of the latter type with suitable restrictions on the variance matrix of the true disturbances so that *the* basic model to be considered has the form of a recursive system in which the relationships can be treated one by one and the parameters in them can be estimated by the method of least squares. With all due deference to his opinions I do not believe that we should restrict our vision in this way, although I entirely agree that at particular stages in the development of a subject single-equation methods and least-squares estimation may be justified by practical considerations.

Like many issues that recur in contemporary debates on method, this problem, in its economic aspect, was considered in some detail by Marshall. In Appendix I to his *Principles* he discusses Jevons's criticism of Ricardo and reaches the conclusion that "the greatest objection of all to his formal statement of his central doctrine is that it does not represent supply price, demand price and amount produced as mutually determining one another (subject to certain other conditions), but as determined one by another in a series. It is as though when three balls A, B, and C rest against one another in a bowl, instead of saying that the position of the three mutually determines one another under the action of gravity, he has said that A determines B, and B determines C. Someone else however with equal justice might say that C determines B and B determines A".

Whatever may be true of the individual economic unit over a minute period of time, I believe that the above statement gives a good account of the typical situation actually encountered given the degree of aggregation over economic units and over time with which economists usually work. If this is so then simultaneous equation models are simply a general representation, suitable for estimation and other statistical purposes, of the state of affairs with which theoretical economists believe that they are commonly faced.

From a practical point of view simultaneous equation methods themselves give rise to numerous difficulties. The problems posed by data collection, by considerations of substantive theory and by estimation problems may mean that the adoption of simultaneous equation methods receive from some investigators, like myself, a rather low priority in applied work. But this is quite a different question and one which I shall not try to elaborate further tonight.

In section 3.2 of his paper Professor Wold refers to "some obscurity in the approach of interdependent relations" and asks what applications can be made of models involving relationships of this kind. The answer is, I think, that they may enable us to handle an important feature of the observations under investigation and so enable us to improve our analysis. Marschak's familiar example of the problem of attempting to measure consumers' response to controlled prices when the only available evidence relates to their responses under free market conditions illustrates this point.

This brings me to an entirely different question. I must confess to a good deal of difficulty in understanding Professor Wold's Fourfold Map of Ends and Means in Applied Statistics. It seems to me that the techniques listed in the S E quadrant are in fact likely to appear everywhere else in the table so that the distinctions between the quadrants will become even more blurred

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

than Professor Wold suggests. Thus, for example, the analysis of variance may be employed for what I should classify as description, as when we try to discover whether there is a significant difference between the seasonal means of a time series. The dummy variables for seasons do not give a causal explanation except in a superficial and circumscribed sense, as can be seen by reflecting on seasonal patterns in different parts of the world. An exactly similar remark applies to the regression of mortality on age. Again these techniques arise in the S W quadrant, whereas the branches of substantive theory shown in the S W quadrant are also important in the S E quadrant. In fact investigation in any quadrant involves data, theory and a means of relating them, and any formal scheme which does not recognize this fact explicitly cannot in my opinion get to the heart of the similarities and differences of the various branches of inquiry.

It is good for us all, concerned as we are with our own detailed investigations, to relax now and again and try to see what we are doing against a wide vision of scientific situations and their analysis. Professor Wold has provided such an opportunity this evening and I have great pleasure in proposing a hearty vote of thanks to him for his paper.

Mr. J. DURBIN: It gives me particular pleasure to second this vote of thanks as I had the good fortune to visit Professor Wold in Sweden several years ago and to discuss with him then some of the ideas which he has expounded in his paper. Professor Wold's distinguished contributions to economics and statistics range over a very wide field from the empirical to the theoretical. It is characteristic of him that his book *Demand Analysis*, which sets out to explain the demand for food in Sweden, should at the same time contain a brilliantly lucid exposition of the elements of the theory of stochastic processes. He is therefore unusually well equipped to deal with the subject of his paper.

The basic question raised by the paper is: How far are we justified in using statistical methods based on probability theory for the analysis of non-experimental data? This is one of the eternal questions of statistical inference and we can be grateful to Professor Wold for his illuminating discussion of it. In statistical work we often find ourselves forced to assume that our residuals have convenient stochastic properties, not because we believe that they are produced by anything remotely resembling a chance mechanism but because of lack of knowledge about the processes which generate them. One sometimes wonders what relation the resulting tests of significance have to probability statements based on the frequency interpretation of probability. At times one feels inclined to murmur, "not much". Yet in spite of misgivings, if we are determined to fit a model to the data, to express our uncertainty about the unexplained residuals in probability terms usually seems the most reasonable thing to do.

Professor Wold contrasts the unhappy lot of the non-experimental statistician with that of his more fortunate experimental colleague, who is able to arrange matters so that his unexplained residuals have convenient stochastic properties. My view is that the distinction between experimental and non-experimental situations is not as fundamental as Professor Wold would have us suppose. There are experimental situations in which it is impossible to specify with precision the stochastic properties of the residuals, as for example when the results of the experiment emerge in the form of time series. Similarly, there are non-experimental situations in which standard statistical procedures can be applied quite validly. For example, if we wanted to investigate by means of a sample survey the dependence of household food expenditure upon income and family size we could lay out the sample according to a two-factor factorial design and apply the analysis of variance quite validly in the ordinary way. Thus, although there is undoubtedly a tendency for the unexplained elements to be more under control in experimental than in non-experimental situations, the relation as I see it is by no means, to use Professor Wold's phrase, one of causal dependence.

My own belief is that there is a continuous spectrum ranging from situations on the extreme left, where the stochastic properties of the observations can be specified with complete precision, as, for example, in sampling inspection, to situations on the right, such as those in econometrics, where it is only by a long stretch of the imagination that the use of probability concepts seems justified at all. I was much impressed by the arguments of Sir Ronald Fisher who, in his article in the *Journal* Series B (1955, Vol. 17, p. 69) criticized the use of a simple frequency interpretation of tests of significance in situations other than those on the extreme left of the spectrum. I sometimes wonder whether it would not be better to abandon the use of fixed significance levels in situations of this kind and go back to the use of Karl Pearson's P as a measure of the discrepancy between observation and hypothesis. When Pearson introduced his goodness-of-fit test, the result of the test was assessed by evaluating the probability of getting a value of χ^2 as great as or greater than that observed. The advantage of reintroducing this procedure for all tests would be that in discussing the rationale of the test, one would not feel under as much pressure to inter-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

pret it in frequency terms as with existing procedures. The scientist would merely take the value of P as one factor in the situation in deciding his attitude to the hypothesis under test. One ought to be able to work out a theory of choice of test statistics and of the sensitivity of tests similar to that of Neyman and Pearson for a fixed significance level. I appreciate that decision theory can be given such a broad interpretation that all these approaches can be regarded as decision functions. However, the matter of emphasis is an important one, as Fisher's paper made clear.

Professor Wold is right to emphasize that the statistical methods to be applied to any body of data depend on the stochastic assumptions made. However, I think he goes too far in suggesting that where the correct assumptions are in doubt large sample methods should be used for lack of anything better. Asymptotic properties can be very illuminating but they can also be misleading in small sample situations. Consequently I think it is dangereus to lay down a general principle such as the author's relation (7).

An illustration of how one can be misled by following such a principle too closely is provided by the author's treatment of the regression analysis of time series in the presence of serial correlation. He recommends the application of least-squares methods directly to the raw data on the ground that the resulting estimates of the regression coefficients are consistent, and he suggests that the effects of serial correlation in the residuals should be allowed for merely by making the large sample correction to the variance given by his expression (14). Yet it is well known that if the amount of serial correlation is substantial the resulting estimates of the regression coefficients will be inefficient. In this situation Cochrane and Orcutt suggested that the analysis should be based on the first differences of the observations. I do not think it is generally realized how big an improvement results from this simple device, even if the serial correlation is not very large. For example, I have worked out the effect of applying the device to the simple case of a regression on a single independent variable where the residual and the independent variable are both generated by linear Markoff processes with the same parameter ρ. I find that the ratio of the variance of the regression coefficient in the untransformed analysis to the variance in the transformed analysis is asymptotically $\frac{2(1+\rho^2)}{(3+\rho)(1-\rho)}$. This is greater than unity for $\rho > \frac{1}{3}$ which represents a negligible amount of serial correlation. For $\rho = 0{\cdot}7$, which is by no means high, the ratio is about $2{\cdot}7$. These results suggest that in any regression analysis of positively auto-correlated time series the device can usually be expected to lead to an improvement.

Towards the end of his paper Professor Wold considers the causal pattern of economic behaviour and the effect that this has upon the choice of regression models. He maintains that in the real world there is no such thing as simultaneous interdependence and that a satisfactory economic theory must be based on the unilateral response of individual economic units to changes in their environment. I am prepared to believe that he is right on this point. However, he goes on to argue that the statistical model must reveal an explicit recognition of the causal chain generating the observations. I do not think that this claim is justified. Many econometric analyses are necessarily based on yearly data, whereas the response times between interacting economic units may be very much smaller than a year. In the Stock Exchange, for example, the interactions between buyers and sellers take place at such speed that to all intents and purposes they may be regarded as taking place simultaneously. Consequently, although I am prepared to accept that economic variables are generated by what Professor Wold calls a causal chain system, I see no reason why a macro-economic model for annual data should not take the form of a system of simultaneous regression equations of the kind developed by Haavelmo and the Cowles Commission. I do not believe that the two systems are incompatible. This is not to say that in my view simultaneous regression systems should be used for fitting demand functions. Although I see nothing wrong with these systems as theoretical models I believe they usually fall down for statistical reasons when we try to fit them in practice. That, however, is another story.

The vote of thanks was put to the meeting and carried unanimously.

Professor P. A. P. MORAN: Professor Wold's paper directs our attention back to fundamentals. The necessity of this may be seen from the very long time it has taken statisticians to understand just what regression analysis is and how it differs from correlation analysis.

Professor Wold states that under general conditions of regularity, maximum likelihood estimators are asymptotically unbiased. This, I suppose, means "consistent", as the literal statement is not true. In fact, it is easy to construct simple examples in which the estimator of a parameter, whilst having a distribution which tends to normality in the ordinary manner, has no moments of any order, and thus cannot be even asymptotically unbiased.

I am glad to see path coefficients being used. While their real value is in genetics, they form a very useful shorthand which enables us to clarify our minds in thinking about other problems.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Finally, a remark on the problem of identifiability. It does not seem to have been noticed that there are important cases where an underlying linear relation, although not identifiable however much data we have, can nevertheless be restricted in various ways, for example, by putting inequalities on the coefficients. Thus a linear relation between two variables, both observed with random errors of unknown variance, cannot in general be identified but it may be possible to assert that its slope lies between two identifiable bounds. This is sometimes very useful.

Dr. KLEIN: I think that Professor Wold is justified in being concerned with estimation from non-experimental data, but I would suggest that, in the field of application in which he is most interested, econometrics and social science, some of the techniques of the experimental psychologists could be adopted and certain experimental methods, possibly controlled by some economic variables, could be used to avoid some of the difficulties which arise with the non-experimental data.

Following up some of the points made by Mr. Durbin, I would say that by judicious use of sampling procedures and controlled field surveys in gathering mass observations one can impose certain restrictions on the samples so that more conventional statistical methods can be used. One can either control variables explicitly or control two groups so that they have the same underlying distribution. I think sociologists call this an ex-*post facto* experiment.

Another point put forward by Professor Wold, concerning the use of subject matter theory, I think should be emphasized to include almost all subject matter and knowledge that one can put one's hand on in dealing with the problem. A good knowledge of the institutional structure is important. For example, in a problem closely related to Professor Wold's work, the demand analysis of Henry Schultz, Wold appears to say in his paper that Schultz was correctly following subject matter theory in making the regression of demand on prices. Earlier in the paper he mentioned the work of Karl Fox of the American Department of Agriculture, who says just the opposite, based on some sort of institutional knowledge of agricultural practices in a self-sufficient economy, namely, that meteorological phenomena largely determine crops and supply and that one should make a regression of prices on quantity, instead of quantity on prices, following consistently Professor Wold's arguments about causal chains.

When Professor Wold comes to the question of interdependent systems of equations versus causal chains, I would disagree very much with the emphasis on causal chain mechanisms. I think the other speakers have dealt adequately with this problem, though one might adopt an entirely different view about the applicability of interdependent systems. In any case, I think the best way of looking at the matter is simply to say that causal chain systems are special cases of interdependent systems in the sense that we form a triangular matrix of coefficients of endogenous variables instead of the square matrix used in interdependent systems.

Professor Wold draws a distinction in his paper between a causal chain system in a simple cobweb theory of demand and supply analysis and an interdependent system. I think he is unfair to the interdependent system. He leaves us with the impression that an interdependent system is static—at least the one that he has talked about—that it produces a peculiar sort of distribution, a conditional distribution instead of a total distribution.

With very slight modifications I can construct an interdependent system which is as dynamic as Professor Wold's causal chain system and to which least squares statistical methods would not be applicable. For example Professor Wold had a causal-chain system something like this:

$$\text{(i)}\quad D_t = f(P_t) + u_t$$
$$\text{(ii)}\quad S_t = g(P_{t-1}) + v_t$$
$$\text{(iii)}\quad P_t - P_{t-1} = h(S_t - D_t) + w_t$$

Instead of Wold's interdependent system I would suggest that we would drop the causal chain idea and use:

$$\text{(i)}\quad D_t = f(p_t) + u_t$$
$$\text{(ii')}\quad S_t = g(p_t) + v_t$$
$$\text{(iii)}\quad P_t - P_{t-1} = h(S_t - D_t) + w_t$$

I would use equation (i), drop the lag in equation (ii), which would give me my interdependence, and for (iii) instead of having identity between supply and demand I would simply allow price fluctuations to depend on the difference between supply and demand. It is a dynamic system, and if I

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

wanted to allow some effect for time passed up to the current date and make it even more dynamic I could have price changes depend on historically accumulated differences between supply and demand:

$$(iii') \quad P_t - P_{t-1} = h\left(\sum_{i=-\infty}^{t} (S_i - D_i)\right) + w_t$$

But in any case it is a dynamic interdependent system, and one is under no compulsion to restrict oneself to comparative static systems with interdependence.

The other point I want to emphasize is that there is nothing curious about identities in interdependent systems. We do not need an identity to make a supply-demand model complete. We can use, instead, an adjustment process with a degree of dynamic push to the system, bringing about equilibrium through price fluctuations. Where we have purely accounting identities in interdependent systems, it is interesting that the statistical methods developed for these systems impose the rule that identities play no essential role, and this is as it should be. They are simply definitions, and it does not matter how they are eliminated from the system; they have nothing to do with the statistical problem of estimating parameters, and it is a virtue of the interdependent systems that they do this.

In the analysis of savings and consumption, people are always a little mystified by the fact that the regression of savings on income gives a different correlation coefficient from the regression of consumption on income. Since savings are related to income and consumption by an identity it should not make any difference in an interdependent system which way you normalize the coefficients.

Professor BARTLETT: I should like to add my welcome to Professor Wold. One may note that of the first four speakers, three have been in the field of econometrics; this is not an accident because Professor Wold's interest is in that field, but there might, I think, be a slight danger of the discussion taking either one of two lines, econometrics as such or the more general fundamentals which have been raised. I should rather like to go back to the question of fundamentals.

Professor Wold has raised a very important topic. Professor Barnard said to me before the meeting that it is useful to have a paper of this kind now and again if only to show the younger statisticians what difficulties they may be up against.

While agreeing with some of the previous speakers who suggested that there is not a hard and fast division between experimental and observational data and that there is a continual gradation, I think, nevertheless, Professor Wold is right in stressing that there is a broad division between them. When I had the good fortune to go to Sweden to a conference in statistical psychology, I tried to stress that the specification problem—that is, the sort of assumptions one sets up—depends essentially on one's objective and this can vary; this, of course, would apply to experimental data as well as to observational data. The advantage of the experimental situation is that if the objective is rather more empirical in character one can get along very nicely, as Professor Wold has suggested, because the observations can be arranged in a random and independent manner; but in the observational situation, even with a less ambitious programme, one has still to worry about all the complications in the observations, which will depend on the model or situation underlying them. This complication of dependence is particularly true of all time-series data, or stochastic process data, as it is sometimes referred to. In another paper in Brussels I tried to discuss the complications which arose in analysing stochastic process data in general. One example I gave is perhaps worth mentioning again. When one is studying an epidemic a situation may arise in which there is no exact model and yet one wants to make exact tests. Here, because of the lack of models, one does not know at the moment any valid way of analysing the data.

Professor Wold seems to try to distinguish between causal analysis and stochastic assumptions. Here there is continual gradation: I would not contrast causal assumptions with stochastic assumptions, but would regard causal analysis as sometimes also stochastic in character. This, for example, would apply to epidemic data where some of the consequences are largely the result of what one might term stochastic elements in the general situation.

Professor Wold has referred in his paper to the question whether the statistician needs a thorough-going knowledge of his subject-matter; where observational data are concerned the statistician must be an expert in the field in which he is working (or collaborate with someone who is). In practice this will force him to specialize in one field, or to work in a field with which he is familiar, as Professor Wold says. When I was in the States recently I found the question of what is statistics and what is statistical inference to be a matter of some controversy. A thorough knowledge of the observational field in which a statistician is working is obviously essential. Some

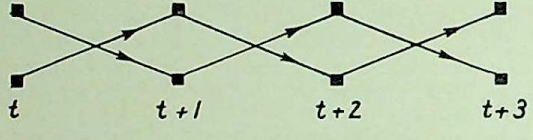
CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

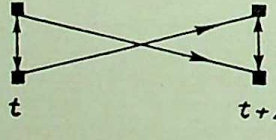
people also seem to demand a *complete* statistical or probability model before any inference can proceed. My own opinion is that that would be a very narrow view to take.

Mr. F. E. A. Briggs: Professor Wold's name is closely linked with causal or recursive regression models, the use of which he has done much to encourage. Undoubtedly there exists a large sphere for the application of such models. More general, however, than the concept of causation is that of association. Causation may be regarded as a particular type of association in which the explanatory variables precede, or occur simultaneously with, the explained variables and causal circles do not exist between those variables occurring smultaneously. If all the variables occur simultaneously then knowledge of some variables will shed information on others provided that the multivariate probability distribution of the variables is known. In the simultaneous equation systems for which estimation procedures have been developed by the Cowles Commission, the joint probability distribution of the endogenous variables is conditioned by the exogenous variables. In such circumstances stochastic circles composed of conditional probability distributions are not inconsistent. Consequently the concept of association permits closed loops while Professor Wold's causal inference does not. Association permits interdependence, whereas causation does not.

If the causal theory is nevertheless accepted and the instantaneous feed-back referred to by Professor Wold is rapid relative to the observation period of the variables, then the observed values of the endogenous variables will be interdependent. This situation is familiar in economics where many of the data are presented on an annual basis, and has given rise to the regression methods developed by Haavelmo and the Cowles Commission. The fact that interdependence of the observed values of the endogenous variables may result in such situations may easily be derived algebraically, but can also be seen in terms of Professor Tinbergen's arrow schemes. If the observation period is twice that of the causal relation then it can be seen that the arrow scheme



collapses into the interdependent form



Consequently in many practical situations causal inference requires resource to the more general inferences of the theory of association. Included in this category of situations are those forms of demand analysis in which total expenditure is taken as the determining variable.

Professor Wold has emphasized the importance of randomization of the uncontrolled factors in simple linear regression. A similar importance attaches to randomization in interdependent regression systems. Such systems come outside the scope of causal inference, and the requirements of randomization cannot be met consistently if each relation is treated by the method of causal inference, in which interdependent endogenous variables become explanatory variables of the causal relations. Consider for example the interdependent regression model.

$$A\,x + B\,z = u$$

in which x are endogenous variables, z are exogenous variables, and u represent uncontrolled factors randomized with respect to the exogenous variables, and with zero mean. The covariance matrix of x and u is given by

$$E\,x\,u' = A^{-1}\,E\,u\,u'$$

indicating that the uncontrolled factors are not in general randomized with respect to those endogenous variables which become explanatory variables in a causal interpretation of the model. For in an interdependent model the matrix A^{-1} is not triangular. In a recursive model A^{-1} is

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

triangular and randomization is achieved if the uncontrolled factors entering each regression equation are independent of each other.

Failure to take into consideration this lack of randomization in a causal interpretation of an interdependent model results in inconsistency, bias and inefficiency of the estimates obtained from such an interpretation. Such an interpretation leads to an underestimate of the uncontrolled variation resulting in turn in an underestimate of the standard errors of the estimates. Adjustment can, however, be made to the least squares method for this lack of randomization and the methods of the Cowles Commission may be derived in this way. Provided that the randomization of the uncontrolled factors with respect to the exogenous variables is maintained, the resulting estimates are consistent and under certain conditions asymptotically efficient. This randomization can only be achieved by including the correct endogenous variables in the regression system, illustrating the close correspondence which exists between model building and experimental design.

Causation is part of a more general concept of association. Many practical problems in causal inference cannot be solved without resource to the theory of association and any attempt to apply a causal interpretation in such circumstances is misleading.

Professor M. G. KENDALL: May I add my appreciation to that of the previous speakers at hearing this paper and at Professor Wold's presence. He and I are such old friends that he will forgive me if I pass over the courtesies and proceed at once to the discourtesies.

Robert Browning has a quatrain:

> There's a great text in Galatians,
> Once you trip on it, entails
> Twenty-nine distinct damnations,
> One sure, if another fails.

If we substitute an appropriate reference to errors in equations for the text in Galatians we get a fair representation of the present position in econometrics. I will not deal with all 29 of the difficulties but I would like to mention four of them. I take an econometric example, not because I am not interested in fundamentals, but because I think of them in concrete patterns.

Let us consider a simple expression which purports to express the demand in terms of price and income (or the logarithms of those quantities):

$$d_t = \alpha p_t + \beta i_t \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The first difficulty is to know what that means. "Demand" can be used either to denote what people would like to buy, or to mean what they actually do buy, and it makes a great deal of difference in the way the equation is treated which precise significance is put on demand variables. That is the first difficulty.

The second difficulty is to know what is meant when such an equation is written down. Any mathematician knows what it means, but the economist does not necessarily mean the same thing; and here I join with previous speakers in differing from Professor Wold about this question of time lag. Strictly speaking, all these equations should have time elements in them, but we often suppose that for the period of observation the adjustment of response to stimulus is nearly instantaneous. One danger of this is that equations are forced into giving identities which do not really exist. It is still not clear in a great deal of econometric writing whether equations of this kind are dynamic equations of motion, or equations of steady state.

The third difficulty, if we have managed to make up our minds on the first two, is whether the above equation is to be regarded as a functional equation or not. I think the answer is that it must be so. If that is correct, attempts to estimate α and β by regression analysis may be completely off the point. We might write down demand as a regression of price and income, as follows:

$$d_t = \alpha P_t + \beta I_t + \varepsilon_t \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The difference between the (1) and (2) is obvious.

The point comes out particularly when we are using equations of this kind in econometric work. If it is desired to use equation (2) as a whole as a predictor, the theory of regression may be useful; but what is done more often is to take one of the coefficients, for example β, and to use it as an elasticity of demand on income. This may be quite wrong.

The fourth point is an elaboration of this last point. There are what are called multicollinearities

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(which should be called collinearities) where linear or approximately linear relations exist between the independent variables in a regression system. I announce without proof the following therem: that if only enough variables are used in a regression system every system is almost collinear, and consequently the matrix determining the coefficients is ill-conditioned. No coefficient by itself in a regression equation has then any considerable validity.

I adduce these points as illustrations of the sort of difficulty we are in and the kind of difficulties to which Professor Wold has directed our attention. I am sure he would be the first to admit that his paper is not the last word on the subject, scarcely perhaps the first, and we are indebted to him for raising these interesting problems.

Professor WOLD subsequently wrote as follows:

I should like to express my gratitude to the Society for accepting my paper for this meeting. I feel somewhat audacious in choosing such a broad topic for discussion, and, further, I am aware of skating on thin ice in some passages of the paper. I am, therefore, all the more interested in the comments made, and appreciative of the course the discussion has taken; also, I must add, relieved, for the knowledge of theoretical and applied statistics embodied in this meeting is impressively authoritative.

Professor Bartlett's remarks confirm my experience that behind the streamlined façade of text-books and research reports the fundamental questions about ends and means in statistics are being debated in many circles. This is only as it should be, a sign of vigour of our science, and it may be added that the debate is closely linked up with the modern development of statistics, for with the ever-expanding range of fruitful applications the basic principles of statistical inference have come to the fore in the general theory of knowledge. Professor Bartlett further gives support to the integrated picture of statistical methods displayed in my fourfold map. If we consider the following general pattern for statistical inference,

$$\text{sample} \xrightarrow[\text{degree of confidence } P \text{ ; loss function } W]{\text{inference}} \text{population} \qquad . \quad . \quad . \quad \text{(A)}$$

where the arrow symbolizes the inference from the sample or samples to the hypothetical population or populations, and the loss function serves as a brief reference to Wald's theory of decision functions, it can be said that the current text-books on statistical method arrange their material in the logical order imposed by the mathematics that comes into play in scheme (A). In one chapter, for example, the testing of the difference between a theoretical and an observed mean is treated together with the testing of the difference between two observed means, and in another chapter correlation and regression methods are brought together. The type cases of applied work are usually discussed by way of numerical illustrations. The fourfold map provides the frame for a systematic treatment of the principles of applied statistics, but its purpose is only to supplement the scheme (A), not to replace it.

This last remark is a reply to Professor Stone's final comment. I am afraid I do not understand him when he refers to the fourfold map as a formal scheme. On the contrary, the map places the mathematical structure of scheme (A) in a perspective which links it up with the concrete ends and means of applied statistics. Incidentally, the concrete significance of the map is strengthened by the interpretation it gives of the broad path NW $\rightarrow$ SE $\rightarrow$ SW of applied statistics as being a matter of scientific evolution. Further I am puzzled by one or two of Stone's introductory general remarks. My illustrations in section 3 refer to experimental *v.* non-experimental situations at various stages of scientific development, and how is the stage of development to manifest itself if not in a penetrating and well-integrated theory? To repeat my line of argument in a simile, an integrated theory is like an iceberg, and the small part visible above the water is the hypothesis that is specified in the analysis of the given data ; a self-contained experiment makes explicit use only of the visible part, but with observational data the hypothesis testing will often draw material from the part of the iceberg that is submerged. When Stone reflects that a sophisticated theory is of great use to econometricians, saying at the same time that the theory has been carried too far in some directions, I appreciate this bit of irony.

Professor Bartlett points out that the assumptions behind a causal model may be stochastic or nonstochastic according to circumstances. This is very true, and illustrations (i)–(ii) in section 3.3 were actually chosen with regard to this distinction, the Pareto-Slutsky theory of consumer choice being non-stochastic, while Porter's migration model is stochastic. It had escaped my attention that the terms causal *v.* stochastic assumptions introduced in 2.3 are unfortunate from

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

this point of view (they were chosen with regard to the simple situation in the *t*-test, regression analysis and the like), and I realize now that it is much better to speak instead of causal *v. residual* assumptions. This term may be substituted throughout my paper.

Several speakers comment upon a feature which is stressed in my paper, namely, that there are no hard and fast lines of demarcation in the fourfold map. Thus, when Mr. Durbin says that with observational data the degree of confidence P is only one factor of many to be taken into account when judging the hypothesis under test, his argument is entirely in line with my remarks in 1.3 (iii) and with the gradual shifts of emphasis in formulae (5)–(7) and (9). I find it difficult, however, to follow him when he suggests that the frequency interpretation of P might be abandoned. Barred from such interpretation—even if it were highly fictitious—the number P would obtain the status of a magic symbol, to be worded not as "probability" but otherwise, say as "Pearson number" or the like.

Durbin refers to his striking illustration of the gain in efficiency when applying the difference method in a case of strongly autocorrelated residuals. The knowledge of such cases is one of the reasons why I have indicated in (7) the small sample theory as an ideal. As regards the variate difference method it may be remarked, on the other hand, that it throws away relevant information when the long-term variations in the series under analysis are not ruled by the same interrelation as the short-term variations (cf. Wold-Juréen, 1952, Ch. 15.3).

Professor Kendall's comments give me a pleasant feeling that we are tuned to the same wave-length. The problems of regression analysis go to the core of the matter in causal analysis by statistical methods, and in practical work it is necessary to take some standpoint or other in the controversial questions at issue. I am looking forward to studying the work he has mentioned on the four specified questions, and I take this opportunity to refer to a review of regression methods where some of these points have been discussed (Wold-Juréen, 1952, esp. Ch. 2; for the first question, see Chs. 3.1 and 7.7 and footnote 4 in Ch. 3.5; for the second, Chs. 1.2–3 and 2.2; for the third, Ch. 2.4; for the fourth, Ch. 2.6). As to the fourth point I suppose that Kendall's theorem needs some restriction relative to the number of variables and observations, for as now announced the theorem becomes trivial if there are more variables than observations, while it becomes untrue if the variables are independent and we are dealing with an infinite number of observations.

Similarly, I am pleased to state general agreement with Dr. Moran's remarks. The existence of moments is one of the requisite conditions of regularity.

Most of the speakers have commented upon simultaneous systems, and in particular upon (a) the rationale of causal chain systems *v.* systems of interdependent relations, and (b) the general scope of the two types of system. My paper deals with (a) from the point of view that in applied statistics it is necessary to specify the operational purpose of the inference, and that once this has been done the estimation problems become a purely technical matter. Everybody seems to agree that the causal chain systems are quite clear-cut in this respect; for example, if the demand relation

$$d_t = D(p_t, \mu_{t-1}) + z_t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(B)}$$

is involved in such a system it can be given the same causal interpretation as an ordinary demand function, thus indicating how demand d will change with lagged income μ if price p and other influencing factors are constant, and similarly how demand d will change with price if lagged income and the disturbance factors are constant. Hence, in the linear case *the parameters of* (B) *can be interpreted as demand elasticities in Marshall's sense;* on the usual non-correlation assumptions about the residual z they can be estimated by least squares regression, and the resulting estimates can be used for practical purposes in accordance with Marshall's definition of elasticities. In particular they can be used for making demand forecasts. As regards the systems of interdependent relations I am not at all convinced that the situation is as clear as would appear from Professor Stone's and Dr. Klein's comments. In the interpretation of such systems I am influenced by C. Christ,* and I gather from Mr. Briggs's comments that he is on the same line; for example, if demand d_t and price p_t are jointly dependent upon lagged income μ_{t-1}, this means that d_t and p_t are related by a probability distribution where μ_{t-1} enters as a conditional variable. The joint conditional distribution can be specified in terms of two interdependent relations, but these do not allow the same direct causal interpretation as (B), and in particular their parameters cannot be interpreted as ordinary demand elasticities (cf. the italicized sentence above). This in itself is

* National Bureau of Economic Research, New York (1951): Conference on business cycles, p. 35. See also Wold-Juréen (1952), Exercise IV, 31.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

not obscure. What I find obscure is that the estimates derived from interdependent relations are still labelled as demand elasticities, as in the works of Bergstrom (1955) and Hildreth and Jarrett (1955). As pointed out in section 3.2 (v) of my paper, it would seem that such "elasticities" are subject to an upward bias, a suggestion which has received no comment from the supporters of interdependent systems here present. The forecast procedures add to the ambiguity; thus K. Fox (1953) derives demand elasticities from interdependent relations and applies them just as ordinary elasticities, whereas Hildreth and Jarrett treat the prediction problem along the same lines as Christ.

I welcome Dr. Klein's remark about the reference to Fox's forecasts given in section 1.4 (ii); my neglect to mention the difference in method is due to a mistake in my notes. Since Fox makes use of interdependent relations, his forecasts raise the question to what extent they are affected by the above-mentioned bias in the elasticity estimates in this approach. Further I appreciate Klein's reference to the notion of *ex post facto* experiment. The illustrations in 3.1 borrowed from Cochran and Welander are typical instances, and I repeat that a main difference from controlled experiments lies in the absence of randomization. This difference can hardly be stressed too much as it seems to be a frequent misconception that the random selection of sampling units in an *ex post* experiment can serve as a surrogate for the *ex ante* randomization of controlled factors in a genuine experiment. This last remark has no bearing upon Dr. Klein, of course, nor upon Mr. Briggs, who speaks of randomization in a somewhat different sense.

Coming finally to question (b), the scope of causal chains *v.* interdependent relations, the situation has a parallel in the system

$$\frac{dx_i}{dt} = F_i(x_1, \ldots, x_n); \quad i = 1, \ldots, n \qquad . \quad . \quad . \quad . \quad (C)$$

of differential equations. In philosophical literature model (C) is sometimes called "the standard form of causal relations". Such a label is deceptive, for although (C) is of very general scope the notion of cause and effect is not associated with a particular mode of mathematical treatment. For the same reason it would be inappropriate if a system of the causal chain or the interdependent relations type were set forth as "the" basic model of dynamic economics. Since Professor Stone suggests that I have placed the causal chain systems on such a pedestal, here I must make a disclaimer (cf. the reference to Tustin in 3.2 (v); see also Wold-Juréen, 1952, p. 70).

In model (C) the variables x_i at instant $t + dt$ are expressed in terms of the variables x_i at instant t. It may be said that (C) is less general, formally, than models of the causal chain (briefly CC) or the interdependent relations (IR) type; in fact, if we regard dt not as an infinitesimal but as a finite time period, (C) becomes a system of the same type as the *reduced form* of a CC or an IR model. In other words, the difference between CC and IR systems lies in properties that cannot be illustrated by systems of type (C); the difference lies on another plane. This is a noteworthy conclusion, I think, and in particular we see that the causal circles which I have discussed in section 3.2 (v) are an obscurity that can never occur in systems (C). The argument is relevant to comments by Stone, Briggs, Durbin and Klein. Thus in the classic problem of three bodies the motion is ruled by a system of type (C), and here the three bodies, say A, B, C, influence each other mutually, and it cannot be said that the interaction is consecutive, say A acting upon B and B upon C. This well-known fact, however, has nothing to do with the question of causal circles in the interpretation of IR systems. Similarly, Stone's reference to Marshall's three balls in a bowl is beside the point with regard to the causal circles, but here the situation is static, and I have preferred to choose instead the dynamic problem of three bodies for illustration. In reply to Briggs and Durbin, I share their view that CC systems have not been designed to cope with the case when the chain makes more than one circuit within the unit period of observation, and I may add that this argument, to the best of my knowledge, first came up in a paper by R. Bentzel and myself (1946, section 16–17). Further, in reply to Briggs and Klein, I share the view that IR systems are more general than CC systems, but the crucial question is whether the operational meaning of the relations involved remains the same when passing from CC to IR systems. I understand from Briggs's comments that he agrees that interdependent relations do *not* have the same causal meaning as the relations in a causal chain system. I am not quite sure, however, whether he then agrees with my conclusion in 3.2 (v) regarding elasticities derived on the basis of interdependent relations, like those reported by Bergstrom, Hildreth and Jarrett, namely, that either they are not elasticities in the ordinary sense or they are biased estimates of such elasticities. Except for the possible difference in this respect, I find myself in complete agreement with Briggs's comments on CC and IR systems.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

As a result of the ballot taken during the meeting, the candidates named below were elected Fellows of the Society:

James Bacon
Richard Michael Batkin
Madan Lal Bishnoi
Janet Thelma Blatcher
Bryce Arthur Murray Cottrell
Hilda Margaret Davies
Paul Duncan-Jones
Iorwerth Lloyd Evans
Michael Gent
Alan Geoffrey Horsnail
Ronald Owen Hughes
Alan Jeffrey
Clara Rose Kerry
Colin Lingwood Mallows
Sidney Reuben Maxwell
Alfred George Simms
John Bernard Spencer
Roebert Lyle Stearman
Guy Woodrow

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Some Aspects of the Structure of Indian Industry

By Vera Anstey and Russi J. Taraporevala

SUMMARY

The following article analyses some of the data provided by the Indian Census of Industries, 1950, with the object of providing a tentative sketch of the structure of India's 28 largest industries.

Reference is made to their scale, size-groups and the characteristics of these groups, relative economic significance, degree of mechanization, levels of salaries and wages, average earnings, and labour productivity.

An attempt at a comparison of the characteristics and structures of these industries is made in tables which give their relative ranking with reference to eleven selected criteria.

It is concluded that although in certain limited directions comparisons can be drawn from the Census data, much more information—quantitative and qualitative—needs to be collected and analysed before any accurate account can be given of the structure of Indian Industry.

Introduction

Few attempts have so far been made to analyse statistically the nature of industrial structure in India. It is, indeed, only since an Industrial Census was instituted that anything approaching comprehensive and reliable figures have been available, even for the larger and more important industries.

The tables attempt to apply to the numerical information available for Indian industries methods of classification and comparison that have been found helpful in other countries. The bases of classification for various purposes have been necessarily limited by the form in which the material is presented in the Industrial Census, whilst the interpretations that can be suggested are still more severely limited by the acknowledged defects and incompleteness of the data. For this reason the comments upon each table are mainly explanatory, whilst the use made of the tables, and any conclusions based thereon, are extremely tentative and subject to amendment, or correction, if or when fuller and more accurate data become available. Indeed, one of the most valuable conclusions that emerges may well prove to be the desirability of securing fuller, and in some cases, different data.

The preliminary results of the fifth Census of Manufactures, 1950, have been published as they have come available in the *Monthly Statistics of the Production of Selected Industries of India*, and it is from these monthly issues that our primary data have been gleaned. These results are "provisional"; coverage is incomplete, and in some cases returns are subject to correction. In the final report considerable improvement in the quality and coverage is expected.

The Census covered some 7,300 factories (or establishments) in 17 states of the Indian Union, in most cases on a statutory basis.*

The Census is confined to factories within the purview of the Factories Act, 1934: and to those employing 20 or more workers on any day, and using power. "Organised manufacturing industry" in India has been classified into 63 groups, following the United Nations industrial classification so far as is practicable. But in India (and other less developed countries) it is extremely difficult to secure a satisfactory classification, and the best that could be done was to classify each factory according to its principal product (value basis). The Census is confined to

* The excluded areas are the Andaman and Nicobar Islands, Bhopal, Bilaspur, Cooch Bihar, Coorg, Hyderabad, Jamnu and Kashmir, Kutch, Madhya Barat, Manipur, Mysore, Saurashtra and Tripura. It is stated that "Except for Hyderabad and Mysore, the number of factories in the States not covered by the Census is negligible."

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the first 29 groups, which include the 28 major industrial activities of India and coal-mining (the latter is excluded from our tables).

"Fixed capital" refers to the position on December 31st, 1949, and indicates original cost, plus cost of improvements made, minus deterioration in value since purchase. The value of land and of buildings is shown separately.

The figure for "fixed capital" is only a very rough guide for our purposes. In general it must always be remembered that the use of accounting data for purposes of economic analysis may give rise to a number of difficulties and misconceptions, especially if used for comparisons, if only because it is probable that accounting procedures vary in certan respects between industries.

"Total cost" is defined as (*a*) wages and salaries, plus (*b*) materials consumed. It excludes depreciation, interest, insurance, taxes, advertising charges, etc. Hence statistics of total production costs are not available nor can profits be calculated.

The "value of materials consumed" is based on what is spent to bring the materials to the factory (excluding goods merely bought and re-sold), including fuel and electricity consumed (unless produced in the factory); and payment for work done on materials given out by the factory. The "value of products" is taken at ex-factory level for those sold, plus the value of those held in stock (whether fully or semi-manufactured), and receipts for work done, or sales (of, for instance, fuel or power) to others, less payments made to outside (e.g. transport) firms. Hence "value added by manufacture", representing the net result of manufacturing activity, is defined as "the total value of products minus the *total* value of materials consumed". But as in estimating "total cost" and "material cost" depreciation has been ignored, the value shown as added by manufacture is a little more than it should be.

In the Census, estimates were made of depreciation for both total and material costs, but these figures were omitted in actually using cost figures, partly because the estimates were considered to be unreliable, and partly because their magnitude was small; hence their omission made little difference to the final results, so that it did not seem worth while to make the complicated computations which their inclusion would have necessitated.

"Employees" are divided into wage and salary earners, the former including only those engaged in manufacturing processes or in cleaning machinery and premises. (Those employed in distributive activities, e.g. sale and advertising, are entirely excluded.) "Wages" cover money payments (including allowances and overtime, etc.) only.

Non-cash benefits have not been estimated separately for wage and salary earners. The total bill for wages, salaries and benefits in kind has simply been divided by the total number of employees of all types in order to obtain average (annual) earnings.

"Man-hour" statistics collected relate only to wage-earners, but have been estimated for workers employed through contractors and for salary earners. "Labour productivity" is obtained by dividing the value added by manufacture by the input of man-hours of all employees. This therefore represents (at 1950 prices) "the net result of industrial activity per man-hour" in each industry.

As the following tables are derived from the monthly statistics (instead of from a completed Census), they include information issued at different periods. This is not always consistent, or apparent inconsistencies are unexplained. For instance the July, 1952, issue (Table 1, p. 4) gives the total number of factories from which returns were received as 6,323, whereas the September, 1952, issue (Table 1, p. 5) gives a total of 6,002 factories only (excluding closed factories). Were 321 factories closed? Or did some factories fail to give the requisite data for this table? We do not know. Again, "rounding" sometimes is, and sometimes is not practised, giving rise to minor discrepancies (other examples could be quoted). For these reasons—as well as because of the acknowledged deficiencies of the Census itself (some of which have been noted above)—we wish to emphasize the tentative nature of our analyses, and to stress that each table is representative only of the data provided in relation to its particular features.

In particular, it should always be borne in mind that the whole Census information is on a "factory" or "establishment" basis (no distinction being drawn between the two). No data are given on the basis of a firm, which may of course own a number of factories (or establishments). This is true of the figures provided by most other countries.

We therefore claim no more than that the following tables present a first approximation to a picture of the structure of India's main manufacturing industries, so far as this can be gauged

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

from the provisional monthly statistics. We hope to evoke criticisms and suggestions and, better still, further research.

TABLE 1

Range of Employment

Industry	*Small Factories Number of Employees Between*	*Medium Factories Number of Employees Between*	*Large Factories Number of Employees Between*	*Number of Factories (alll sizes) at work*
1. Rice Milling	20 and 50	51 and 100	101 and 196	1,248
2. Fruit and Vegetable Processing	,,	,,	101– 364	21
3. Starch	,,	,,	101– 179	21
4. Soap	,,	,,	101– 1,102	47
5. Wheat Flour	20– 100	101– 200	201– 490	64
6. Biscuit Making	,,	,,	201– 469	73
7. Distilleries and Breweries	,,	,,	201– 252	40
8. Vegetable Oils	,,	,,	201– 866	897
9. Paints and Varnishes	,,	,,	201– 662	39
10. Tanning	,,	,,	201– 783	76
11. Plywood and Tea-chests	,,	,,	201– 618	27
12. Sewing Machines	,,	,,	201– 1,480	6
13. Electric Lamps	,,	,,	201– 505	8
14. Electric Fans	,,	,,	201– 2,003	27
15. General Engineering and Electrical Engineering	,,	,,	201– 3,328	1,639
16. Woollen Textiles	20– 150	151– 300	301– 2,654	39
17. Iron and Steel	,,	,,	301–23,033	133
18. Bicycles	,,	,,	301– 1,070	21
19. Glass and Glassware	20– 200	201– 400	401– 726	96
20. Matches	,,	,,	401– 2,027	37
21. Ceramics	20– 250	251– 500	501– 2,150	52
22. Chemicals	,,	,,	501– 2,110	225
23. Aluminium, Copper and Brass	,,	,,	501– 2,299	223
24. Paper and Paper-Board	20– 300	301– 600	601– 3,743	36
25. Sugar	20– 500	501–1,000	1,001– 1,618	278
26. Cement	,,	,,	1,001– 1,991	14
27. Cotton Textiles	20–2,000	2,001–4,000	4,001– 8,274	508
28. Jute Textiles	20–3,000	3,001–6,000	6,001–10,607	107
			Total	5,992

Source: *Monthly Statistics of the Production of Selected Industries of India*, September, 1952, pp. 6, 7, 8.

N.B.—This table gives the Census definition of "small", "medium" and "large" factories—on the basis of numbers employed—for each industry. The Census was confined to factories "which employed 20 or more workers on any of the days of the year 1950 and carried on the manufacturing process with the aid of power". (*Monthly Statistics*, December, 1951, p. 4.)

For the purposes of this paper factories are divided into three groups—small, medium, and large—based on the range of employment, as shown in Table 1. Each industry has its own definition of size groups, and the maximum of the range of employment in the large factories denotes the employment size of the largest factory in the industry. The Census provides no data for any other definition of size group. The Table shows that in every industry there are factories which vary considerably in size, often from the pigmy to the giant.

Table 2 shows the distribution of the total number of factories between the three size groups, as defined by the Census. In the last column the average number of salaried and wage workers per factory in each industry is given. It can be seen that except in the cement and glass and glassware industries a large majority of the total number of factories (actually between 61 per cent. and 93 per cent.) belong to the small size group. In only three industries (glass and glassware 34 per cent., cement 28 per cent., and jute textiles 29 per cent.) are over 20 per cent. of the factories in the medium size group; in only four industries (electric lamps 25 per cent., woollen textiles 26 per cent., paper and paper board 28 per cent., and cement 36 per cent.) are over 20 per

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 2

Size-grouping of Factories

	Small Factories		*Medium Factories*		*Large Factories*		
	Number of Factories	*Percentage of the Total Number of Factories in the Industry*	*Number of Factories*	*Percentage of the Total Number of Factories in the Industry*	*Number of Factories*	*Percentage of the Total Number of Factories in the Industry*	*Average Number of Salaried and Wage Workers per Factory*
1. Rice Milling	1,093	88	137	11	18	1	39·4
2. Fruit and Vegetable Processing	14	67	3	14	4	19	80·9
3. Starch	16	76	3	14	2	10	47·6
4. Soap	37	79	4	8	6	13	114·9
5. Wheat Flour	47	73	11	18	6	9	82·8
6. Biscuit Making	59	81	7	9	7	10	87·6
7. Distilleries and Breweries	28	70	6	15	6	15	112·5
8. Vegetable Oils	830	93	52	5	15	2	56·8
9. Paints and Varnishes	31	80	3	7	5	13	115·3
10. Tanning	57	75	11	14	8	11	109·2
11. Plywood and Teachests	19	70	4	15	4	15	122·2
12. Sewing Machines	4	67	1	16	1	17	333·3
13. Electric Lamps	5	63	1	12	2	25	200·0
14. Electric Fans	18	67	4	14	5	19	207·4
15. General Engineering and Electrical Engineering	1,403	86	139	8	97	6	88·1
16. Woollen Textiles	25	64	4	10	10	26	384·6
17. Iron and Steel	117	88	4	3	12	9	575·9
18. Bicycles	17	81	3	14	1	5	128·6
19. Glass and Glassware	55	57	32	34	9	9	221·9
20. Matches	27	73	3	8	7	19	335·1
21. Ceramics	41	79	1	2	10	19	332·6
22. Chemicals	206	92	10	4	9	4	148·9
23. Aluminium; Copper and Brass	201	90	16	7	6	3	111·2
24. Paper and Paperboard	22	61	4	11	10	28	597·2
25. Sugar	214	77	51	18	13	5	347·8
26. Cement	5	36	4	28	5	36	1035·7
27. Cotton Textiles	385	76	84	16	39	8	1176·3
28. Jute Textiles	71	66	31	29	5	5	2868·2

Source: *Monthly Statistics*, September, 1952, pp. 6, 7, 8. The order of industries follows that of Table 1. The average numbers of workers per factory are obtained from the figures given in Tables 1 and 4.

cent. of the factories in the large-size group. It is interesting to note that a preponderance of small-scale factories is also to be found in the advanced countries of the world, such as the United Kingdom and the United States. According to the U.S. Census of manufactures, 1939, no less than 73 per cent. of all industrial units employed not more than 20 persons. In the U.K., the Census of Production, 1935, estimated that no less than 73 per cent. of the total number of industrial units employed less than 11 workers, accounting for 9·3 per cent. of all industrial workers. Moreover 59 per cent. of the establishments employing 10 or more workers, employed not more than 50 workers. The largest number of workers were, however, employed by the groups of establishments employing between 200 and 499 workers (Florence, 1943, pp. 23 and 34).

But the method of grouping which gives these results and others to be noted later appears completely arbitrary. For instance, the percentage of the total number of factories classified as "small" is greater for certain industries—notably the chemicals (92 per cent. "small") and the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

aluminium, copper and brass industry (88 per cent. "small") for which all factories employing up to 250 persons are classed as "small", than for other industries, such as rice-milling (only 88 per cent. "small") and starch (only 76 per cent. "small"), although in the latter industries no factory employs more than 200 persons, even in the "large" group. Unless carefully interpreted, this may give an erroneous impression.

In other words, although this grouping may show the distribution of various sized factories within each industry in relation to the particular structure of the industry concerned, it fails to show clearly the distribution of size groups between industries.

To take another instance, 93 per cent. of the factories in the vegetable oils industry, and 92 per cent. of those in the chemicals industry are classed as "small". But for the former, "small factories" are defined as those employing up to 100 persons, whereas in the latter, "small" factories are defined as those employing up to 250 persons. Had the same definition been adopted for both industries, a very different result would have been obtained. If the definition for all industries of "small" factories had been those employing up to 2,000 persons (as for cotton textiles), the percentage of small to total factories would have been 100 for no less than 17 industries. (This, in its turn, would not have been very illuminating.)

It can only be concluded that the Census grouping is based on judgment decisions on what sizes, and spread of sizes, are typical (or representative) for each industry. These judgment decisions may, or may not, be useful and realistic. But in any case, interpretations based on figures resulting from this type of classification must attempt specifically to avoid the bias or even distortion that would otherwise arise. It may, indeed, be suggested that it might be better to collect figures relating to size or scale on the same basis for all industries, in such a manner that they could be regrouped at will for any particular purpose. In the absence of such data we have used the average number of employees per factory in each industry as some indication of the relative size of establishments as measured by their labour force.

Though the number of small factories is much greater than that of the medium or large, the economic significance of the small factories compared with that of the two latter categories varies greatly between industries, when criteria other than mere numbers are used. In many cases their significance is much less than that of "medium" and/or "large" factories. Economic significance can be measured in a number of ways. We have adopted three yardsticks, i.e. plant and machinery (in value terms), number of employees (wage earners and salaried); and the value of the product. Table 3 gives the percentages of these three items accounted for by factories in each of the three size groups out of the totals for each industry. The Table shows that the various industries can be grouped in relation to the relative importance of the factories in each size group.

In rice milling the small factories, with 83 per cent. of plant and machinery, 71 per cent. of employees, and 75 per cent. of the value of products, show a clear dominance over the medium and large factories—the last size group being least important. In the next group of four industries (Nos. 2–5) small factories account for some 49 per cent. of our three criteria of economic significance. But there are substantial differences within this group. In the vegetable oils industry the small factories stand out, accounting for some 59 per cent. of the three criteria, whilst the medium and large factories are of minor, but not dissimilar, importance. In the jute-textile industry, small and medium factories are of almost equal importance (48 per cent. and 40 per cent. respectively), with the balance slightly in favour of the former: large factories are relatively insignificant. In the chemicals and distilleries industries the small factories constitute 48 per cent. and 43 per cent. respectively of each industry but are closely followed in significance by the large units (30 per cent. and 33 per cent.), with the medium-sized units of least importance.

Next come three industries, i.e. wheat flour, aluminium, copper and brass, and cotton textiles, where no particular size-group predominates, each accounting for approximately one-third of the industry. Then we have the sugar and glass, and glassware industries, in which the medium-sized factories are most important and account for some 50 per cent. of each industry.

We have next a group of eight industries (Nos. 11–18) in which the large factories account for between 46 per cent. and 68 per cent. of each industry. In each case, except that of cement (in which the small and medium factories are of almost equal importance), the small factories come second in importance, medium-sized factories being least significant. Then come six industries (Nos. 14–24) in which large factories account for over 75 per cent. of each, followed (excepting the electrical fans industry for which small and medium factories are of almost equal importance) in

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

significance by the small and then the medium factories. The case of iron and steel is especially interesting, as twelve factories (or 9 per cent. of the total) account for some 93 per cent. of the industry.

TABLE 3

Percentage Distribution of Plant and Machinery, Employees and Product between Large, Medium and Small Factories in each Industry.

Industry	*Small Factories*			*Medium Factories*			*Large Factories*		
	Value of Plant and Machinery %	*Number of Employees* %	*Value of Product* %	*Value of Plant and Machinery* %	*Number of Employees* %	*Value of Product* %	*Value of Plant and Machinery* %	*Number of Employees* %	*Value of Product* %
1. Rice Milling	83	71	75	14	23	21	3	6	4
2. Vegetable Oils	52	67	57	31	18	20	17	15	23
3. Jute Textiles	48	47	48	44	40	39	8	13	13
4. Chemicals	48	50	48	13	15	17	39	35	35
5. Distilleries and Breweries	53	40	38	21	24	27	26	36	35
6. Wheat Flour	43	37	37	24	30	44	33	33	19
7. Aluminium, Copper and Brass	35	46	34	32	25	36	33	29	30
8. Cotton Textiles	49	33	34	30	36	37	21	31	29
9. Sugar	39	29	33	44	50	49	17	21	18
10. Glass and Glassware	23	24	22	55	48	53	22	28	25
11. General Engineering and Electrical Engineering	45	40	33	17	16	14	38	44	53
12. Tanning	34	32	33	14	21	27	52	47	40
13. Biscuit Making	30	34	27	17	20	21	53	46	52
14. Plywood and Tea-chests	29	32	27	13	18	16	58	50	57
15. Fruit and Vegetable Processing	49	29	24	11	13	9	40	58	67
16. Starch	11	29	33	6	23	16	83	48	51
17. Cement	33	16	19	20	24	24	47	60	57
18. Ceramics	45	26	17	2	2	2	53	72	81
19. Paper and Paper-board	25	9	5	7	9	6	68	82	89
20. Matches	25	13	4	8	6	3	67	81	93
21. Woollen Textiles	20	12	11	8	7	5	72	81	84
22. Soap	13	21	8	7	6	2	80	73	90
23. Electric Fans	16	12	8	16	13	10	68	75	82
24. Iron and Steel	4	8	5	0·3	2	1	95·7	90	94
25. Sewing Machines	6	10	4	—	—	—	94	90	96
26. Electric Lamps	11	16	16	—	—	—	89	84	84
27. Paints and Varnishes	70	35	30	13	10	7	17	55	63
28. Bicycles	16	26	11	54	28	23	30	46	66

Source: Compiled from data given in *Monthly Statistics*, September, 1952, pp. 6, 7, 8. The Census is by law prohibited from disclosing data for individual establishments. Hence in some cases where the number of factories is very small (in practice less than 3) separate figures for certain states and hence for certain size-groups are withheld. For the sewing machines and electric lamps industries, owing to the small number of both medium and large factories, these are taken together in this Table. Where the number of factories is less than 3 for only one size-group (for instance the Starch Industry has only 2 "large" factories) the missing figures have been found by deduction.

N.B.—The order of industries in this table is roughly based on the relative significance (as judged by our 3 criteria) of the three size-groups. First come those in which small factories predominate (this predominance being greatest in the first industry, and so on); then those where no size-group predominates; next those where the medium-size, and thereafter those where the large-size predominates, this predominance being greatest for Tea and Steel, the last of this group. Finally come the "odd" cases, i.e. sewing machines and electric lamps (medium and large-scale combined); and paints and varnishes and bicycles, for which the predominant scale varies greatly as between our three criteria.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Finally we have four industries which need special explanation. Separate figures for medium and large factories are not available for the sewing machine and electric lamps industries. This is because the number of factories in the large-scale group is so small that separate figures would reveal the position of particular concerns a result that is prohibited by a "secrecy" clause in the legislation providing for the Census. In practice, where there are less than three factories in any group separate figures are not given, but taken together these two size groups account for some 93 per cent. and 86 per cent. respectively of each industry.

The paints and varnishes industry shows very wide differences in the percentage of plant and machinery, employees, and value of products respectively within each size group. For instance, small factories account for 70 per cent. of the plant and machinery, but only 30 per cent. of the value of products and 35 per cent. of employees; whereas large factories account for 63 per cent. of the value of products, but only 17 per cent. of plant and machinery. Possibly this may be explained by the heterogeneous qualities of the products, and the corresponding specialization of factories of particular size-groups in particular products—characteristics shared by many other industries—and the fact that large factories specialize upon mass production of inexpensive products, whereas the smaller factories produce small quantities of various high-grade goods. Nevertheless, it remains astonishing that in the size-group with most (70 per cent.) plant and machinery the value of products should be relatively so low, whilst in the size-group with only 17 per cent. of the industry's plant and machinery the value of products should be so high (63 per cent.).

The bicycle industry also shows variations of the same type, though these are less extreme. Small factories account for only 11 per cent. of the industry's value of products, but 26 per cent. of the number of employees, whereas large factories account for 66 per cent. of the industry's value of output, but only 30 per cent. of the plant and machinery.

On the whole there appears to be an extremely close relationship in most industries (there are only a few exceptions) between the share of plant and machinery, the share of employees, and the share of the value of products, in each industry's total, falling into each size group. The other main exceptions, in addition to paints and varnishes and bicycles, are the fruit and vegetables, starch, cement and match industries. In the last case, for instance, although small factories account for 25 per cent. of the plant and machinery, they account for only 4 per cent. of the value of products. It might have been expected, on *a priori* grounds, that the size group with the largest share of plant and machinery would also have the largest share of value of products, and a smaller share of the number of employees. But apparently this is not so, nor does any clear pattern or relationship emerge.

Examination of Table 3 shows that in different industries different size-groups of factories have varying economic significance. These may be accounted for by various factors, e.g. technical, managerial, financial and marketing, which affect the structure of each industry. It might be conjectured that the "optimum" size—a concept very dear to the heart of some economists—might lie in the size-group of greatest importance in the particular industry in question. But this would assume that competition is sufficiently strong to force factories towards the optimum size, so that the larger part of the industry would come within the optimum size-group (as measured by our three yardsticks), whilst factories outside this size-group (and hence, by hypothesis, *less efficient*) are eliminated. We doubt whether this assumption is justified in an economy like India, where many industries are in their infancy and often enjoy partial monopoly advantages. It is impossible to discover from statistics whether there is any one optimum size of factory even at a particular moment of time (in this case, when the Census data were gathered). But our analysis might be useful as a pointer to different size-groups of factories, which are most important in particular industries, although the fact that the classification of an "industry" is based on the nature of the predominant product, irrespective of other types of product may render generalization of this type somewhat illusory. For instance a "bicycle" factory may be substantially engaged in manufacturing particular "parts", or in assembly work, which may favour a different structure from that needed by a factory which only manufactures complete bicycles.

Table 4 sets forth the relative economic significance of the various industries, as judged by the number of wage and salary earners employed (which is used for ranking purposes in the Table), and by the value of plant and machinery employed in each industry. In some cases the ranking would be notably different if based on the value of plant and machinery. For instance, the sugar and rice milling industries employ relatively little plant and machinery, whilst the iron and steel,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 4

Relative Economic Significance of Industries

Industry	*Total Number of Wage and Salary Earners per Industry* (1,000s)	*Percentage of Wage and Salary Earners in all Industries*	*Total Value of Plant and Machinery in the Industry* (Rs. 1,000)	*Percentage of Plant and Machinery Employed in all Industries*
1. Cotton Textiles	597·6	38·9	369,180	26·9
2. Jute Textiles	306·9	20·0	159,985	11·6
3. General Engineering and Electrical Engineering	144·3	9·4	139,189	10·1
4. Sugar	96·7	6·3	90,461	6·6
5. Iron and Steel	76·6	5·0	168,925	12·3
6. Vegetable Oils	51·0	3·3	102,646	7·5
7. Rice Milling	49·1	3·1	26,129	1·9
8. Chemicals	33·5	2·2	63,231	4·6
9. Aluminium, Copper and Brass	24·8	1·6	39,464	2·9
10. Paper and Paper Board	21·5	1·4	61,432	4·5
11. Glass and Glassware	21·3	1·4	15,023	1·1
12. Ceramics	17·3	1·1	7,194	0·5
13. Woollen Textiles	15·0	1·1	8,049	0·6
14. Cement	14·5	1·0	48,824	3·6
15. Matches	12·4	0·8	3,435	0·2
16. Tanning	8·3	0·5	3,201	0·2
17. Biscuit-making	6·4	0·4	6,684	0·5
18. Electric Fans	5·6	0·4	5,865	0·4
19. Soap	5·4	0·4	9,879	0·7
20. Wheat Flour	5·3	0·3	10,381	0·8
21. Paints and Varnishes	4·5	0·3	5,445	0·4
22. Distilleries and Breweries	4·5	0·3	8,738	0·6
23. Plywood and Tea-chests	3·3	0·2	5,177	0·4
24. Bicycles	2·7	0·2	4,438	0·3
25. Sewing Machines	2·0	0·1	2,196	0·2
26. Fruit and Vegetable Processing	1·7	0·1	2,040	0·2
27. Electric Lamps	1·6	0·1	4,607	0·3
28. Starch	1·0	0·1	1,355	0·1
	1,534·8	100·0	1,373,173	100·0

Source: Estimated from *Monthly Statistics*, various issues, containing details for each industry. The summary tables in the July, 1952, and September, 1952, issues differ slightly from the totals secured from the detailed tables partly due to rounding, and partly because for the Sugar Industry the detailed figures (April, 1952) give season and off season employees (used here), whereas the summary tables give only season employees. This does not, however, affect the ranking of the industries.

cement, soap, wheat flour and distilleries and breweries industries employ relatively much plant and machinery.

The cotton textiles industry emerges as the most important industry according to both these indices. It is followed by the jute textiles, general engineering and electrical engineering, sugar, and iron and steel industries—although the ranking of these industries differs according to whether the "employees" or the "plant and machinery" basis is used. These industries, including the cotton textiles industry, employ no less than 79·6 per cent. of the total number of employees covered by the Census, and 67·5 per cent. of the total value of plant and machinery used. They should, therefore, be considered as the main industries of India to-day—the backbone of her industrial structure. The next seven industries (numbers 6–12) employ 14·1 per cent. of the total number of employees and 23·0 per cent. of the total value of plant and machinery used in all industries. The remaining sixteen industries thus account for only 6·3 per cent. of the total employees, and for 9·5 per cent. of the total value of plant and machinery used in all industries.

The Census figures give only a very rough guide to the degree of mechanization of any industry, partly because of the particular method by which plant and machinery is valued—i.e. on the basis of original cost, plus cost of improvements, minus deterioration in value since purchase—and partly because in general the use of accounting procedure for our purposes imposes serious limita-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 5

Degree of Mechanization

Industry	*Value of Plant and Machinery per Worker (Wage-earner) In Rupees*	*Industry*	*Value of Plant and Machinery per Worker (Wage-earner) In Rupees*
1. Cement	3,842	15. Fruit and Vegetable Processing	1,333
2. Electric Lamps	3,428	16. Electric Fans	1,262
3. Paper and Paper Board	3,175	17. Biscuit Making	1,243
4. Iron and Steel	2,713	18. Sewing Machines	1,234
5. Vegetable Oils	2,489	19. Sugar	1,163
6. Wheat Flour	2,419	20. General Engineering and Electrical Engineering	1,125
7. Chemicals	2,409	21. Glass and Glassware	759
8. Distilleries and Breweries	2,340	22. Cotton Textiles	662
9. Soap	2,133	23. Rice Milling	652
10. Aluminium, Copper and Brass	1,849	24. Woollen Textiles	598
11. Bicycles	1,862	25. Jute Textiles	544
12. Plywood and Tea-chests	1,861	26. Ceramics	469
13. Starch	1,526	27. Tanning	440
14. Paints and Varnishes	1,524	28. Matches	310

Source: Compiled from *Monthly Statistics*. The number of wage-earners or "operatives" (i.e. excluding salary earners) as defined for the purpose of the Census, and the value of plant and machinery, are given in the details for each industry, scattered throughout the issues. (For the latter, see also Table 4 above.)

tions on the accuracy and relevance of the results. The Census method of valuing plant and machinery introduces a purely arbitrary element which varies with the age of the plant and machinery, thus tending to exaggerate the capital intensity of factories with newer plant and machinery, compared with those with older plant and machinery.

Table 5 attempts to depict the extent of mechanization, as measured by the value of plant and machinery per wage-earner, or operative. On the basis adopted the match industry appears to be the least mechanised—a very surprising result as in western countries the capital intensity has been claimed to be large for this industry. The three textile industries (jute, woollen and cotton) have a very similar degree of mechanization (between Rs. 544 and Rs. 662 per operative), which is rather low. Rice-milling, which is essentially a very small-scale industry has, however, only the sixth lowest extent of mechanization amongst all the industries, and appears to be more mechanized than the jute and woollen textile industries. The four most mechanized industries are cement, electric lamps, paper and paper-board and iron and steel (in that order).

In general, eight industries employ less than Rs. 1,000 worth of plant and machinery per operative, eleven employ between Rs. 1,000 and Rs. 2,000, six between Rs. 2,000 and Rs. 3,000, and three over Rs. 3,000.

Table 6 shows the average annual cash earnings of salary-earners and wage-earners, as defined for the purpose of the Census (see above, "Introduction"). In general terms the "wage-earners" are operatives, whilst those employed in a supervisory or clerical capacity are deemed to be "salary earners". The table reveals the following features:

(1) In every industry, the average earnings of the salaried personnel are greater than those of operatives, the ratio of salaries to wages varying, as between industries, from 1·3 to 3·4. The ratio is between 1·3 and 1·9 for five industries, employing a total of some 3,500 salaried persons and 153,400 operatives; it is between 2·0 and 2·9 for twenty industries, employing some 119,200 salaried persons and 165,400 operatives; and between 3·0 and 3·4 for three industries employing 500 salaried persons and 3,600 operatives. For all industries taken together the average ratio is 2·3; this means that a great majority of employees are in industries where this ratio is near the average.

Nevertheless, although the average earnings of salaried persons are invariably greater than those of wage-earners in the same industry, there are a number of cases in which the average

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 6

Levels of Salaries and Wages

Industry	*Average Annual Salaries (Rupees)*	*Average Annual Wages (Rupees)*	*Average Salaries—Divided by Average Wages*
1. Soap	3,046	1,149	2·7
2. Electric Lamps	3,017	912	3·3
3. Paints and Varnishes	2,761	993	2·8
4. Iron and Steel	2,655	1,484	1·8
5. Matches	2,572	1,011	2·5
6. Woollen Textiles	2,543	920	2·8
7. Paper and Paper-Board	2,402	921	2·6
8. Cotton Textiles	2,285	1,101	2·1
9. Bicycles	2,279	1,298	1·8
10. Distilleries and Breweries	2,237	824	2·7
11. Fruit and Vegetable Processing	2,170	630	3·4
12. Chemicals	2,103	930	2·3
13. General Engineering and Electrical Engineering	2,102	955	2.2
14. Sewing Machines	2,038	1,589	1·3
15. Jute Textiles	1,989	769	2·6
16. Cement	1,946	925	2·1
17. Electric Fans	1,938	839	2·3
18. Aluminium, Copper and Brass	1,870	941	2·0
19. Starch	1,783	601	3·0
20. Wheat Flour	1,693	870	1·9
21. Plywood and Tea-chests	1,619	597	2·7
22. Biscuit Making	1,619	709	2·3
23. Ceramics	1,614	745	2·2
24. Glass and Glassware	1,554	589	2·6
25. Tanning	1,431	714	2·0
26. Vegetable Oils	1,270	529	2·4
27. Sugar	1,035	631	1·6
28. Rice Milling	645	266	2·4
Average, All Industries	2,008	873	2·3

Source: Estimated from *Monthly Statistics*, various issues, containing details for each Industry.

earnings of operatives in some particular industry are greater than those of salaried persons in one or more other industries. For instance, the average earnings of salaried persons in rice milling are less than those of wage-earners in all but seven of the twenty-eight industries under consideration. At the other end of the scale, the average earnings of operatives in the sewing machine industry are greater than those of salaried persons in five of the other industries; in the iron and steel industry greater than those of salaried persons in four other industries; in the bicycle industry, greater than those of salaried persons in three other industries; and in the soap and cotton textile industry greater than those of salaried persons in two of the other industries.

(2) The average annual cash-earnings of salaried persons in all 28 industries is Rs. 2,008, but varies greatly between industries; i.e. from Rs. 645 in rice-milling to Rs. 3,046 in the soap industry. The following table summarizes the variations in average "salaries" between industries.

Average Salary	*Number of Industries*	*Number of Salary-earners*	
Under Rs. 999	1	9,054	
Rs. 1,000–Rs. 1,499	3	29,851	
Rs. 1,500–Rs. 1,749	5	6,344	
Rs. 1,750–Rs. 1,999	5	19,568	91,568
Rs. 2,000–Rs. 2,249	5	29,288	
Rs. 2,250–Rs. 2,499	3	42,339	
Rs. 2,500–Rs. 2,749	3	17,388	
Rs. 2,750–Rs. 2,999	1	1,010	
Rs. 3,000 and over	2	1,080	
Total	28	155,922	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

It may be noted that over 50 per cent. of the industries pay average salaries of between Rs. 1,500 and Rs. 2,249; whilst almost 60 per cent. of the salary-earners are in industries paying average salaries of between Rs. 1,750 and Rs. 2,499.

(3) The average annual cash earnings of all operatives is Rs. 873, but here also there are wide variations between industries: i.e. from Rs. 266 in rice milling, to Rs. 1,589 in the sewing machine industry. The following table summarises the position:

Average Wage	*Number of Industries*	*Number of Operatives*
Under Rs. 499	1	40,078
Rs. 500–Rs. 599	3	63,834
Rs. 600–Rs. 699	3	80,201
Rs. 700–Rs. 799	4	321,929
Rs. 800–Rs. 899	3	12,675
Rs. 900–Rs. 999	8	221,719
Rs. 1,000–Rs. 1,099	1	11,025
Rs. 1,100–Rs. 1,199	2	562,533
Rs. 1,200 and over	3	66,437
Total	28	1,380,431

(4) The lowest average salaries and wages are found in the rice milling industry, followed by the sugar, vegetable oils and tanning industries. At the other end of the scale the soap industry pays the highest average salary, whilst its average wage comes about midway in the list. The iron and steel industry pays the second highest average wage and the fourth highest average salary. Hence it and the soap industry are the "best paid" industries.

(5) An attempt was made to correlate the average salaries and wages paid by particular industries in order to discover whether or not high wages went with high salaries. There was found to be a high coefficient of correlation between salaries and wages for the 14 industries paying average salaries of less than Rs. 2,000; a very low coefficient between salaries and wages for the 14 industries paying average salaries of over Rs. 2,000, and a medium coefficient for the whole group of 28 industries. But for various technical reasons it was considered inappropriate to apply, or at least to rely upon, this method. It was therefore decided to fall back upon the "ranking" of industries according to various measurements and criteria, in the hope that some interesting relationships might emerge, and this is done in Table 8. But it may be noted that the ratio between salaries and wages in the lower salaried group of 14 is only 2·26, whilst that in the higher salaried group of 14 is 2·36, and that average wages in the lower salaried group are almost 68 per cent. of average wages in the higher salaried group, whilst average salaries in the former group are only 46 per cent. of those in the latter group. It is thus clear that where salaries range high wages also tend to be higher (there are only two industries in the higher salaried group paying average wages of less than Rs. 900, whereas in the lower salaried group only two industries pay average wages over Rs. 900), but nevertheless tend to lag in relation to salaries.

The Census also gives average annual earnings and indices of labour productivity not only for each industry but also for size-groups within each industry. "Labour productivity" has been measured by dividing the "value added by manufacture" by the aggregate input of man-hours by all employees. (For definitions of "value added by manufacture" and "input of man-hours" see "Introduction" above). The census authorities claim that this "represents in terms of 1950 prices the net result of industrial activity per man-hour in each of the industries for which the census was carried out".

Table 7A gives these figures and indices, whilst Tables 7B, C and D are derived from Table 7A.

Table 7B shows that the highest average annual earnings are obtained in "large" factories in 18 industries (65 per cent. of the total), whilst in "small" factories the highest average earnings are obtained in only 3 industries (10·7 per cent.).

From Table 7C it is interesting to find that in 20 industries (70 per cent. of the total) the "large" factories have the highest labour productivity although in one of these cases first place was shared with "small" factories; in 8 industries the medium-sized factories led, although in one of these they shared first-place with "small" factories. "Small" factories did not stand above "large" and "medium" factories in respect of labour productivity in any industry, although they did

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 7A

Earnings and Labour Productivity Related to Size-Groups of Factories.

Industry	Average Annual Earnings (Rupees)				Labour Productivity			
	Total	*Small*	*Medium*	*Large*	*Total*	*Small*	*Medium*	*Large*
1. Rice Milling	360	343	376	410	0·2	0·3	0·2	0·3
2. Fruit and Vegetable Processing	840	779	535	932	1·0	0·5	0·4	1·6
3. Starch	786	326	482	1,418	0·9	1·0	0·02	1·4
4. Soap	1,538	876	866	1,794	2·8	0·6	0·6	3·7
5. Wheat Flour	1,068	1,040	1,188	991	1·2	0·9	1·7	1·1
6. Biscuit-making	908	789	855	1,018	1·5	1·0	1·1	2·1
7. Distilleries and Breweries	1,102	881	1,194	1,303	0·9	1·0	1·0	0·8
8. Vegetable Oils	685	501	836	1,379	0·7	0·6	1·0	1·1
9. Paints and Varnishes	1,421	1,143	1,059	1,643	2·0	1·3	1·2	2·6
10. Tanning	813	668	879	880	0·8	0·8	0·3	1·0
11. Plywood and Tea-chests	858	754	828	940	0·9	0·5	0·7	1·1
12. Sewing Machines	1,772	797	961	1,949	1·0	0·4	0·9	1·1
13. Electric Lamps	1,300	1,096	2,709	1,026	1·3	1·3	2·6	0·9
14. Electric Fans	1,093	934	1,273	1,092	1·9	0·7	1·1	2·2
15. General Engineering and Electrical Engineering	1,144	860	1,149	1,382	0·7	0·6	0·6	1·0
16. Woollen Textiles	1,119	689	599	1,221	1·2	0·4	1·2	1·4
17. Iron and Steel	1,774	903	965	1,862	1·4	0·4	0·8	1·5
18. Bicycles	1,470	898	1,043	2,090	1·5	0·5	1·4	2·0
19. Glass and Glassware	669	610	705	663	0·4	0·4	0·5	0·3
20. Matches	1,333	397	611	1,527	1·2	0·2	0·3	1·4
21. Ceramics	883	658	609	971	0·5	0·3	0·2	0·5
22. Chemicals	1,224	1,210	1,494	1,132	1·9	1·5	1·7	2·6
23. Aluminium, Copper and Brass	1,099	910	1,322	1,204	0·7	0·5	1·6	0·1
24. Paper and Paperboard	1,124	728	794	1,202	1·0	0·6	0·6	1·1
25. Sugar	735	828	727	631	1·5	1·4	1·5	1·4
26. Cement	1,125	1,331	1,160	1,055	1·4	1·3	1·4	1·5
27. Cotton Textiles	1,191	1,084	1,251	1,234	0·7	0·7	0·8	0·7
28. Jute Textiles	833	858	803	836	0·5	0·5	0·7	0·6
All Industries	1,057	890	1,029	1,269	0·8	0·6	0·8	1·0

Source: *Monthly Statistics*, October, 1952, p.g. *et seq.* Average annual earnings have been obtained by dividing the total of wages, salaries and non-cash benefits by the total number of employees of all types.

TABLE 7B

Pattern of Earnings in Relation to Size-Group

Sequence of Earnings

	Highest	*Second Highest*	*Lowest*	*Number of Industries*	
Size-Group of Factory	Large	Medium	Small	13	18
	Large	Small	Medium	5	
	Medium	Large	Small	4	7
	Medium	Small	Large	3	
	Small	Medium	Large	2	3
	Small	Large	Medium	1	

TABLE 7C

Pattern of Labour Productivity in Relation to Size-Group

Sequence of Productivity	*Highest*	*Second Highest*	*Lowest*	*Number of Industries*	
Size-Group of Factory	Large	Medium	Small	14	20
	Large	Small	Medium	6	
	Medium	Large	Small	4	8
	Medium	Small	Large	4	
	Small	Large	Medium	0	0
	Small	Medium	Large	0	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 7D

Links between Sequence of Average Annual Earnings and of Labour Productivity

Extent of Correlation	*Number of Industries*
Full concurrence of sequence	15
Concurrence of highest columns only . .	7
,, second columns only . .	2
,, lowest columns only . . .	2
No concurrence	2

Source: Derived from Table 7A.

share first place with "large" factories in one industry, and with "medium" factories in another industry.

Table 7D shows that there is a marked concurrence between earnings and labour productivity in relation to the size of factories in a majority of industries, and that there are only two industries in which no degree of concurrence exists.

Rice milling (0·2), followed by glass and glassware (0·4), ceramics (0·5) and jute textiles (0·5) had the lowest labour-productivity; whilst soap (2·8), paints and varnishes (2·0), electric fans (1·9), and chemicals (1·9) led the list.

These results show that, in general, "large" factories give better opportunities for higher earnings than "medium" and "small" factories, and support the prevalent view that "large" factories (at least up to a point, not yet generally attained in India) secure greater technical efficiency, but nothing can be said about profitability on this basis alone.

Tables 8A and 8B attempt to give a picture of the main characteristics of the factories in the 28 industries under discussion, so far as can be judged by the census data. Table 8A gives the list of industries, and 11 headings under which the ranking of each of these industries is entered, The ranking number of each industry is thus entered under the appropriate heading. For instance rice milling stands second as regards number of factories, 7th as regards number of employees, 11th as regards value of plant and machinery, 23rd as regards plant and machinery per head, 13th as regards ratio of salaries to wages, and last as regards the remaining 5 items. Table 8B rearranges the same material so that it can be seen at a glance approximately how each industry stands with respect to ranking under each of the 11 headings. For this purpose the 28 ranked industries are divided into four quartiles (7 industries in each quartile), and under each quartile are entered the numbers of the "items" which fall within that quartile. For instance, engineering ranks amongst the top quartile of industries for "items" 2 (number of factories), 3 (numbers of employees), and 5 (total value of plant and machinery); amongst the second quartile for "items" 7 (average salaries) and 8 (average wages); and amongst the third quartile for all the rest, except for items 4 and 10, which fall into the bottom quartile. Item 10 appears (with a query) in both the last quartiles, because engineering is bracketed equal with several other industries on the margin between the two last quartiles.

The Census definitions of most of the headings in Table 8A have already been given (see Introduction and comments on Tables 6 and 7). Thus "employees" (col. 3) include wage- and salary-earners; "value of plant and machinery (col. 5) is based on original cost, plus cost of improvements, minus the value by which plant and machinery has deteriorated; the value of plant and machinery divided by the number of operatives (i.e. salary-earners not included) accounts for col. 6; col. 7, 8 and 9 are taken from Table 6; col. 10, "labour productivity" is obtained by dividing the value added by manufacture, by the input of man-hours of all employees, and is held to represent "the net result of industrial activity per man-hour." Col. 11, which we have called "operational efficiency" has not previously appeared in this article. It is the obverse of an estimate of "total cost expressed as percentage of value or products", and hence can be taken as a rough measure of the surplus of "value of products" over material and labour costs. Col. 12, "value added by manufacture expressed as a percentage of investment in plant and machinery" is also new to this article.

As explained in our Introduction, the figure for "value added by manufacture expressed as a percentage of investment in plant and machinery" is rather higher than it ought to be, because

TABLE 8A

Comparative Ranking of Industries.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

TABLE 8A

Comparative Ranking of Industries.

Key to Industries	*Number of Factories*	*Number of Employees*	*Average Number of Employees per Factory*	*Value of Plant and Machinery*	*Value of Plant and Machinery per Worker*	*Average Annual Salaries per Head*	*Average Annual Wages per Head*	*Ratio of Salaries to Wages*	*Labour Productivity*	*Operational Efficiency*	*Value Added by Manufacture as % of Investment in Plant and Machinery*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. General Engineering and Electrical Engineering	1	3	22	4	20	13	8	19	22	18	16
2. Rice Milling	2	7	28	11	23	28	28	13	28	28 (—)	28
3. Vegetable Oils	3	6	26	5	5	26	27	14	21	27	27
4. Cotton Textiles	4	1	2	1	22	8	6	21	24	24	10
5. Sugar	5	4	7	6	19	27	22	27	7	15	11
6. Chemicals	6	8	14	7	7	12	10	17	4	2	12
7. Aluminium, Copper and Brass	7	9	20	10	10	18	9	23	23	25	24
8. Iron and Steel	8	5	5	2	4	4	2	26	8	10	19
9. Jute Textiles	9	2	1	3	25	15	18	10	26	17	8
10. Glass and Glassware	10	11	11	12	21	24	26	9	27	21	17
11. Tanning	11	16	21	25	27	25	20	22	20	23	3
12. Biscuit making	12	17	23	18	17	22	21	15	5	5	7
13. Wheat Flour	13	20	24	13	6	20	15	24	11	26	20
14. Ceramics	14	12	10	17	26	23	19	18	25	19	9
15. Soap	15	19	18	14	9	1	4	8	1	16	6
16. Distilleries and Breweries	16	22	19	15	8	10	17	7	18	14	23
17. Paints and Varnishes	17	21	17	20	14	3	7	5	2	7	4
18. Woollen Textiles	18	13	6	16	24	6	13	4	12	12	2
19. Matches	19	15	8	24	28	5	5	12	13	9	1
20. Paper and Paper Board	20	10	4	8	3	7	12	11	16	8	26
21. Plywood and Tea-chests	21	23	16	21	12	21	25	6	19	13	18
22. Electric Fans	22	18	12	19	16	17	16	16	3	1	5
23. Fruit and Vegetable Processing	23	26	25	27	15	11	23	1	14	11	14
24. Starch	24	28	27	28	13	19	24	3	17	22	21
25. Bicycles	25	24	15	23	11	9	3	25	6	4	13
26. Cement	26	14	3	9	1	16	11	20	9	3	22
27. Electric Lamps	27	27	13	22	2	2	14	2	10	6	25
28. Sewing Machines	28	25	9	26	18	14	1	28	15	20	15

Table 8A gives the ranking of the industries listed in column (1) for each item given in columns (2) to (12).

The industries in column (1) are listed according to the number of factories in each industry. Hence the ranking order in column (2) corresponds with that of the list of industries. The figures under each column from (2) to (12) represent the ranking order of the corresponding industries in column (1). Thus taking the first industry, "General Engineering and Electrical Engineering" we find that it has more factories than any other industry; comes 3rd for "number of employees", 22nd for "average number of employees per factory"; 4th as regards "value of plant and machinery", and so on. Alternatively, taking for instance column (7) "Average Annual Salaries per head", we find that Soap ranks first in this respect, Electric Lamps second, Paints and Varnishes third, and so on.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 8B

Summary of Industrial Ranking

Industries	*Top* 25%	*Second* 25%	*Third* 25%	*Bottom* 25%
1. General Engineering and Electrical Engineering	2, 3, 5	7, 8	6, 9, 10?, 11, 12	4, 10?
2. Rice Milling	2, 3	5, 9	—	4, 6, 7, 8, 10, 11, 12
3. Vegetable Oils	2, 3, 5, 6	9	10?	4, 7, 8, 10?, 11, 12
4. Cotton Textiles	2, 3, 4, 5, 8	7, 12	9, 10?	6, 10?, 11
5. Sugar	2, 3, 4, 5, 10	12	6, 11	7, 8, 9
6. Chemicals	2, 5, 6, 10, 11	3, 4, 7, 8, 12	9	—
7. Aluminium, Copper and Brass	2	3, 5, 6, 8	4, 7, 10?	9, 10?, 11, 12
8. Iron and Steel	3, 4, 5, 6, 7, 8	2, 9, 10	12	9
9. Jute Textiles	3, 4, 5	2, 9, 12	7, 8, 11	6, 10
10. Glass and Glassware	—	2, 3, 4, 5, 9	6, 11, 12	7, 8, 10
11. Tanning	12	2	3, 4, 8, 10	5, 6, 7, 9, 11
12. Biscuit making	10, 11, 12	2	3, 5, 6, 8, 9	4, 7
13. Wheat Flour	6	2, 5, 10	3, 7, 8, 12	4, 9, 11
14. Ceramics	—	2, 3, 4, 12	5, 8, 9, 11	6, 7, 10
15. Soap	7, 8, 9, ? 10, 12	5, 6, 9	2, 3, 4, 11	—
16. Distilleries and Breweries	9?	6, 7, 9?, 11	2, 4, 5, 8, 10	3, 12
17. Paints and Varnishes	7, 8, 9, 10, 11, 12	6	2, 3, 4, 5	—
18. Woollen Textiles	4, 7, 9, 12	3, 8, 10, 11	2, 5	6
19. Matches	7, 8, 12	4, 9, 10?, 11	2, 3, 10?	5, 6
20. Paper and Paper Board	4, 6, 7	3, 5, 8, 9, 10, 11	2, 10?	12
21. Plywood and Tea-chests	9?	6, 9, ? 11	2?, 4, 5, 7, 10, 12	2?, 3, 8
22. Electric Fans	10, 11, 12	4	2?, 3, 5, 6, 7, 8, 9	2?
23. Fruit and Vegetable Processing	9	7, 10?, 11, 12	6, 10?	2, 3, 4, 5,8
24. Starch	9	6	7, 10, 12	2, 3, 4, 5, 8, 11
25. Bicycles	8, 10, 11	6, 7, 12	4	2, 3, 5, 9
26. Cement	4, 6, 11	3, 5, 8, 10	7, 9	2, 12
27. Electric Lamps	6, 7, 9, 11	4, 8, 10	—	2, 3, 5, 12
28. Sewing Machines	8	4, 7, 10?	6, 10?, 11, 12	2, 3, 5, 9

although depreciation was estimated in the Census for total and material costs, it was not included when further calculations were made (such as that now under consideration), partly because the estimates were considered unreliable, and partly because their magnitude was small, and hence their emission made little difference to the final results, so that it did not seem worth while to make the complicated computations which their inclusion would have necessitated. "Value added by manufacture" has therefore been arrived at simply by deducting material costs from value of products.

The object of expressing "value added by manufacture" as a percentage of "investment in plant and machinery" was to secure figures useful within a given industry and as between different industries. It was held that a "more or less definite relation" exists between the "net national results of economic activity in an industry" and the "investment in fixed capital" in that industry. But whereas the investment in plant and machinery is "more or less technologically determined for a given industry of a given size, the value of investment in land and buildings (i.e. the other constituents of fixed capital) depends upon the situation of a factory. Hence the Census omitted investment in land and buildings, in order to secure a more generally applicable figure for comparative purposes. It was added that "it is easy to adapt this figure to the total investment in fixed capital in a given industry as soon as the location of a factory is decided."

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

It is interesting to note that the variation between industries under head 12 is extremely large, i.e. from 1,033·8 per cent. for matches to 47·4 per cent. for rice milling. The second highest figure is 527·4 per cent. for woollen textiles, whilst the second lowest figure is 71·8 per cent. for vegetable oils. Hence the figure for matches is quite exceptional. It has been noted, when commenting upon Table 5, that the match industry seems to be the least mechanized of all. The extremely high value added by manufacture in relation to the investment in plant and machinery bears this out.

Further inquiry has shown (see in particular a series of articles on the "Cottage Match Industry" by S. C. Das Gupta, in the *A.I.C.C. Economic Review*, India, of November 15th, 1953; May 15th, June 15th, and July 1st, 1954) that these characteristics of the match industry in India, contrasted with the match industries in industrialized countries, can be, at least partially, accounted for by the fact that in India (in 1950) only some 7 out of a total of 37 factories utilized fully mechanized methods of production. The rest appear to have used mostly hand processes except for splint chopping and veneer peeling. This is of great interest as, whilst it is alleged that costs are no higher and quality no lower in the "cottage" factories, the others obviously employ far more workers per unit of output.

Certainly Table 8B gives a rather clear, if limited, picture of the structure of the match industry, i.e. a very low ranking for the value of plant and machinery and of plant and machinery per worker; a relatively small total number of factories and employees: a rather high ratio of salaries to wages, labour productivity and operational efficiency; high ranking for salaries and wages, and a position at the top of the list for "value added by manufacture as percentage of investment in plant and machinery". Incidentally, it may be noted here that the figures given for value of plant and machinery (col. 4) are apt to be misleading because, as they ignore replacement value, they tend to give undue weight to the newer industries and to those older industries in which substantial replacement of plant and machinery has already been accomplished.

It was hoped that a comparison of the ranking of industries as given in these tables, might reveal some illuminating relationships, or even hint at certain patterns of industrial structure (within the limits imposed by the choice, definition and content of the 11 headings).

For instance, might it not emerge that a high ranking for value of plant and machinery per worker (col. 6) is frequently linked with high labour productivity (col. 10), and with high "operating efficiency" (col. 11), or with high value added by manufacture (col. 12)? Or possibly, owing to the superabundance of manual labour in India, the relationship might be in the opposite direction. Might not unusually high salaries, and a high ratio of salaries to wages (col. 9) result in improved organization and a consequential high "operating efficiency" (col. 11) and labour productivity (col. 10)?

Actually the results tend to be inconclusive and in a sense disappointing. No clear "patterns" appear. The failure of marked relationships to appear may be accounted for by (*a*) the possibility that the nature, definition, and numerical content of each heading is inappropriate; (*b*) the existence of other factors—e.g. foreign control and management, the date of establishment of the industries, age of the plant and machinery, methods of organisation, and siting of the factories, etc.—which outweigh those included in the headings; (*c*) various degrees of monopoly preventing changes and adaptation that would occur under more competitive conditions. Nevertheless a number of interesting and perhaps suggestive facts—some positive, some negative—can be observed.

For instance, as is only to be expected, cols. 2 (number of factories), 3 (number of employees), and 4 (value of plant and machinery) frequently occur for any one industry in the same quartile. All these (i.e. 2, 3, and 4) occur 12 times in the same quartile; in addition to which 2 and 3 occur together 6 times (making in all 18 links between cols. 2 and 3); 2 and 4 occur together 5 times (making 17 links between 2 and 4 in all); whilst 3 and 4 occur together 5 times (making 17 links between 3 and 4 in all). It may be noted that industries ranking high in cols. 2, 3 and 5 (i.e. the first 5 industries on the list) tend to rank low in other respects. These are mainly rather long-established industries (e.g. engineering, vegetable oils and cotton textiles).

Cols. 7 (salaries) and 8 (wages) occur for any one industry in the same quartile 13 times—supporting what was said in our comments on Table 6.

Cols. 10 (labour productivity), 11 (operational efficiency) and 12 (value added by manufactures, as percentage of investment in plant and machinery) occur together for any one industry in the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

same quartile 9 times. In addition 10 and 11 occur together 7 times (making 16 links on all); and 10 and 12 occur together 3 times (making 12 links in all); and 11 and 12 occur together once (making 10 links in all). This at least suggests that these headings have some meaning.

On the other hand there is little or no linkage between 6 (value of plant and machinery per worker) and any one of the cols. 10, 11 and 12. This may, perhaps, be accounted for by the historic basis of valuation of plant and machinery to which attention has already been drawn in the comments on Table 5.

The industries with the largest number of items in the top quartile are (i) iron and steel (7 items), with only one item (ratio of salaries to wages) in the bottom quartile; (ii) paints and varnishes (6 items), with no items in the bottom quartile; (iii) chemicals and soap, each with 5 items in the top quartile and no item in the bottom quartile; (iv) cotton textiles and sugar each with 5 items in the top quartile: this is mainly because they both have a large number of factories and employees and a high average of employees per factory, whereas both industries have 3 items in the bottom quartile. From the point of view of our "ranking" these may claim to be the leading industries.

The paper and paper board and glass and glass ware industries are the only ones figuring predominantly in the second quartile.

The electric fans, biscuit making, plywood and tea-chests, distilleries and breweries, and engineering industries figure predominantly in the third quartile; whilst the rice milling, vegetable oils, tanning, fruit and vegetable processing and starch industries figure predominantly in the bottom quartile—although for different reasons.

Rice milling ranks in the bottom quartile for no less than 7 out of the 11 cases.

The evidence, so far as it goes, with regard to the importance or effects of the level of salaries and wages, and of the ratio of salaries to wages seems purely negative. The industry with the highest average annual wages (i.e. the sewing machine industry) ranks relatively low for salaries and is at the bottom of the list for "ratio of salaries to wages". On the other hand, the industry with the highest ratio of salaries to wages (i.e. the fruit and vegetable processing industry) pays only the eleventh highest average salary, and ranks 23rd for wages. Both industries tend to rank rather low under most other headings, and have a rather similar series of items in each quartile. The starch industry, which ranks third in the ratio of salaries to wages, has a somewhat similar distribution of ranking as the sewing machine and fruit and vegetable processing industries, whilst all these three, plus the electric lamp industry (ranking second in ratio of salaries to wages) have very low ranking for the number of factories (col. 2), number of employees (col. 3) and value of plant and machinery (col. 5). On the other hand, of the 4 industries with the largest number of items in the top quartile, two (paints and varnishes, and soap) do, whilst two (chemicals and iron and steel) do not, rank in the top quartile for "ratio of salaries to wages". It appears, therefore, that the level of salaries and wages and ratio of salaries to wages is not clearly linked with our ranking in general.

It is not, of course, claimed that this type of analysis has "proved" anything. But it may prove useful in suggesting further lines of inquiry and research, some of which may necessitate descriptive, historical and qualitative, rather than quantitative, surveys or analyses, although in many cases this would necessitate shifting the unit of inquiry from the "factory" to the "firm". Indeed data relating to ownership and management, which must obviously refer to firms rather than factories, and to interconnections between firms (arising, for instance, from the Managing Agent System, interlocking directorates and multiple directorships) are essential if a clear picture is to be gained of the structure of Indian industry. Some interesting data on ownership and control of Indian industries are, for instance, contained in a pamphlet entitled *Who Owns India?* by Asoka Mehta (1950). Detailed comparisons with other countries might also prove fruitful.

We did, indeed, make some attempt to collect information about the extent of foreign ownership and management, and the importance of joint stock organization, in the relevant 28 industries. We secured sufficient data to show that there are wide differences in these respects between the industries; but the information was too incomplete to permit any attempt at ranking, or at correlation or comparison with the data presented in Table 8. Some information secured from the office of the Director of Industrial Statistics is, however, given below in Appendix A.

More illuminating quantitative results could probably be secured if the Industrial Census collected data in a more appropriate form for our purposes (see, for instance, one suggestion in our

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

comments on Table 2). Other statistical authorities could also help if they bore in mind the value of collecting and presenting data relevant to industrial structure and organization. Much more could be done to classify and analyse existing information relating to industrial joint stock companies, whilst it would be of great value if some method could be devised enabling comparisons to be made of relative profitability—or at least of trends in the relative profitability of various industries, even if the data referred only to joint stock companies.

Two other factors that may be of great importance in determining the position and structure of industries are (*a*) the date or period at which a substantial proportion of the factories was established; and (*b*) the siting and relative concentration of industries. Although we have not as yet been able to collect the necessary data one or two comments may be tentatively offered.

It has been shown that India's "most important" industries, accounting for 79·6 per cent. of the total employees in all the 28 industries, are the cotton, jute, general engineering and electrical engineering, sugar, and iron and steel industries. Of these, the cotton, jute and engineering industries are the three oldest of India's "western type" industries whilst the sugar and iron and steel industries are amongst the older of the more recently established industries. All except the iron and steel industry rank high chiefly with respect to the number of factories and employees, and value of plant and machinery, and their whole pattern of ranking is not dissimilar—again except for the iron and steel industry.

This suggests, as might be expected, that the older industries have continued along their original lines, whereas the four industries which we have shown to have the highest "ranking" according to our 11 criteria, are mostly more recently developed industries. The iron and steel industry, however, occurs in both lists.

With respect to siting and concentration, the difficulty is that no criterion suitable for our needs has yet been devised. Professor Balakrishna, basing his calculations upon Professor Sargant Florence's techniques, worked out the coefficients of localization of 14 Indian industries in 1937. (Balakrishna, 1948, pp. 17 and 44). Of these 12 are on our list. Five of them had high coefficients (iron and steel, jute, soap, chemicals and paper) and 7 had medium coefficients (sugar, glass, cement, cotton, woollen, matches, tanneries). None for which estimates were made had low coefficients. A high coefficient shows high concentration. But even if this method were extended to all our 28 industries, for the year 1950, it would not give us what we want, as this method is based on political (i.e. provincial) regions; and it is only for such regions that the Census gives information. What we need is some criterion of the degree to which factories are sited in urban or industrial areas which, on the one hand, provide public facilities and services (such as transport electricity, good water supplies) and have easy access to markets and by-industries, but, on the other hand, are subject to the higher costs associated with such areas. For what we want to know is whether the greater costs in such areas do, or do not, outweigh, for particular industries, the advantages of availability of public utilities, etc.; the relative importance of such factors to particular industries; and the extent to which particular industries are sited to the best advantage. If the Industrial Census could provide data relating to the actual location of factories in particular industries—referring to their siting in large, medium and small towns or villages—a suitable index might be constructed. Data of the various types mentioned (and there may well be many others which we have omitted) might throw most valuable light on some of India's most pressing economic problems—problems which are, moreover, common to many underdeveloped countries. For instance, what types of industrial structure, what degree of mechanization, what scale of plant, what extent of urbanization tend to prevail and tend to favour expansion in India, and in other countries with similar economic conditions and problems?

Appendix A

The following table (p. 80) is taken from information kindly supplied by the office of the Director of Industrial Statistics (1, Council House St., Calcutta 1). The figures are for the year after that to which our other figures refer. Industries are in order of those in which the highest proportion of factories are owned by Joint Stock companies. The percentage of factories owned by Joint Stock Companies is 50 for the glass and glassware industry; 15 industries having a higher percentage, and 12 a lower percentage. Nevertheless 58 per cent. of all the factories were owned by other than Joint Stock companies, because the industries in the latter group include those

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Distribution of Ownership of Factories among Joint Stock Companies and Others in the Year ending December 31st, 1951.

Industries	Number of Registered Factories	Number of Factories Making Returns	Owned by: Joint Stock Companies — Number	Joint Stock Companies — %	Owned by: The Individual Proprietors, etc. — Number	The Individual Proprietors, etc. — %
1. Cement	19	19	19	100	—	—
2. Jute Textiles	109	103	100	97	3	3
3. Electric Lamps	9	9	8	89	1	11
4. Sugar	158	144	122	85	22	15
5. Distilleries and Breweries	55	44	36	82	8	18
6. Paper and Paperboard	43	38	31	82	7	18
7. Electric Fans	34	26	21	81	5	19
8. Chemicals	262	240	186	78	54	22
9. Fruit and Vegetable Processing	25	22	16	73	6	27
10. Cotton Textiles	530	485	347	72	138	28
11. Paints and Varnishes	49	45	32	71	13	29
12. Ceramics	69	58	39	67	19	33
13. Plywood and Tea-chests	42	34	21	62	13	38
14. Bicycles	31	28	15	54	13	46
15. Woollen Textiles	48	43	23	53	20	47
16. Glass and Glassware	127	109	55	50	54	50
17. Tanning	91	83	41	49	42	51
18. Soap	55	47	22	47	25	53
19. Wheat Flour	91	77	35	45	42	55
20. Biscuit Making	93	87	38	44	49	56
21. Starch	55	49	19	39	30	61
22. Matches	46	31	12	39	19	61
23. General Engineering and Electrical Engineering	1,930	1,720	636	37	1,084	63
24. Iron and Steel	153	149	52	35	97	65
25. Sewing Machines	9	9	3	33	6	67
26. Aluminium, Copper and Brass	264	225	72	32	152	68
27. Vegetable Oils (1)	1,157	997	315	32	652	68
28. Rice Milling	1,426	1,336	284	21	1,052	79
Total	6,977(2)	6,257	2,600	42	3,657	58

(1) There are two branches of this industry: (*a*) that excluding edible hydrogenated oils; (*b*) edible hydrogenated oils; but the former is by far the larger, having 1,123 factories as compared with 34 for the latter. Of the 31 of the latter that reported 97 per cent. were owned by Joint Stock Companies.

(2) One factory furnished 4 separate returns for departments engaged in different industries: the number of registered factories is therefore 3 less than the total of this column which amounts to 6,980.

(e.g. general engineering and electrical engineering, rice-milling and vegetable oils) which have the largest number of factories. If it could be assumed that factories owned by Joint Stock companies are normally of "more importance" (whether judged by value of plant and machinery, number of employees, or any other suitable criterion) than the others, our table would show industries in order of the relative importance of the Joint Stock principle: its ranking of industries could then be compared with the rankings in Table 8. But we have no evidence as to whether this is, or is not the case. It may be that in some industries the "more significant" factories are not those owned by Joint Stock companies.

Evidence with regard to foreign ownership is very vague and impossible to quantify. We have been told by the Office of the Economic Adviser to the Government of India that the following industries are "substantially" foreign owned and managed: soap, distilling and brewing, paints and varnishes, electric lamps, general engineering and electrical engineering, matches, chemicals, aluminium, copper and brass, paper and paper board: also that the bicycle industry is "substantially" foreign owned but not foreign managed, whereas the jute industry is "substantially" foreign managed but not foreign owned. We do not know the relative degree of foreign ownership:

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the order given above is quite arbitrary. The other industries are not predominantly foreign owned or managed.

The following information supplied by the same source is also of interest:

Soap. A major share of the capital invested is Indian, but a major share of the present production is contributed by Lever Brothers (India) Ltd., which means that there is considerable idle capacity in the Indian sector. There are also a number of small-scale units owned and managed by Indians (but not through Joint Stock companies).

Bicycles. Manufacture of bicycles has been undertaken jointly with British firms (exact share not known), but the management is predominantly Indian.

Matches. The Swedish Match Company has almost monopolistic control in the organized sector, but some 40 per cent. of the output is contributed by the cottage match industry. It is not clear from these data whether all or part (or none) of the cottage industry enterprises are registered under the Factory Act; if a number (probably all those included in the above table) are registered, this may account for our guess (in our comments on table 8) that in many Indian factories certain processes are performed by hand that in the West are performed by machinery. This may explain the low average value of plant and machinery per worker.

Jute Textiles. A recent estimate is that only 25 per cent. of the capital remains in British hands, although management is still predominantly in British hands.

Tanning. A major share of tanning is done by small-scale units organized as private firms or partnerships. In the organized sectors, however, a substantial share is done by foreign units, e.g. Bata Shoe Company and British India Corporation.

It may be noted that the number of registered factories in each industry given in the above table under the heading "Factories making returns" is not very different from the number given in our Table 1. Usually it is slightly higher. This may be explained by the increase in factories in 1951 compared with 1950, or (more likely) by better reporting.

If we assume, for the sake of argument, that ranking in the above table is comparable with the rankings in Table 8, no obvious correlations or patterns suggest themselves, except that those industries with a small percentage of factories owned by joint stock companies have a rather poor ranking record in Table 8. (It is, of course, not claimed that all the rankings in Table 8 are strictly comparable.) This suggests not so much that organization by joint Stock companies is superior, as that joint stock companies are not attracted by those industries with a relatively poor record in other respects.

Appendix B

Size of Factories, 1951

Since this article was written a welcome addition has been made to the census information relating to the size of factories (Summary of Census of Indian Manufactures, 1947–51. *Supplement to the Indian Trade Journal*, October 9th, 1954, pp. 30, 31.) As these new figures are given for one year only it is to be assumed—although no date is mentioned—that they were collected for the first time in 1951. No analysis of the relation between size and other characteristics is given, so the figures cannot be used for our purposes. It is to be hoped that useful analyses will be included in the full report on the Census for 1951, when it is published.

The new information gives for each industry the distribution of factories according to size, as denoted by the number of persons employed, in the following size-groups: (i) Below 20 (the Census covers only factories which employed 20 or more workers on any day during the preceding year; but some of these factories employed less than 20 persons on an average during the year); (ii) 20–49; (iii) 50–99; (iv) 100–249; (v) 250–499; (vi) 500–999; (vii) 1,000–1,999; (viii) 2,000–4,999; (ix) 5,000 and above.

Bibliography

(1) Government of India: Ministry of Industries and Supply: Directorate of Industrial Statistics: (1949) *First Census of Manufactures*, 1946.

(2) Government of India: Ministry of Commerce and Industry: Directorate of Industrial Statistics: (December, 1951 to November, 1952) *Monthly Statistics of the Production of Selected Industries of India.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(3) Supplement to the *Indian Trade Journal*, October 9th, 1954: *Summary of the Census of Indian Manufactures*, 1947–51.
(4) BALAKRISHNA, R. (1948), *Regional Planning in India*. Bangalore Publishing Co.
(5) DAS GUPTA, S. C. "Cottage Match Industry", *A.I.C.C. of India, Economic Review*, November 15th, 1953; May 15, June 15, July 1, 1954.
(6) FLORENCE, P. SARGANT (1943), *The Logic of British and American Industries*. London, Routledge.
(7) MEHTA, ASOKA (1950), *Who Owns India?* Hyderabad, Chetana Prakashan.

Acknowledgments

Our thanks are due to Professors E. T. Grether, R. Jastran, J. S. Bain, E. Beem, H. Allisson, P. Sargant Florence, and to Mr. P. W. S. Andrews and Mr. Basil Yamey, for having read our manuscript and given us the benefit of their valuable comments and criticisms. Our thanks are also due to Drs. R. K. Nigam and B. B. Ghosh, and Messrs. S. Nanjundar, Harish Chandra, and D. N. Rangnekar for information generously supplied. We are also most grateful for the facilities provided by the Economic Research Division of the London School of Economics, and for assistance in checking the figures generously given by Dr. B. S. Bhir.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Principles of Factor Analysis

By M. G. Kendall and D. N. Lawley

This note is the outcome of a correspondence between the authors following the appearance of the review by M. G. Kendall in this Journal (Series A, Vol. 117, 1954, pp. 482–483) of the Uppsala Symposium on Psychological Factor Analysis, 17th–19th March, 1953. It is hoped that a brief exposition of the main points at issue may help to dispel the confusion which still exists about factor analysis, a confusion which is scarcely surprising considering the lack of clarity of much literature on the subject.

Factor analysis is often confused with principal component analysis, which it formally resembles to some extent. In reality the two methods have entirely different aims. In component analysis a linear and orthogonal transformation is applied to the p variates $x_1, x_2, \ldots, x_p$ to produce a new set of uncorrelated variates $y_1, y_2, \ldots, y_p$. These are chosen such that y_1 has maximum variance, y_2 has maximum variance subject to being uncorrelated with y_1, and so on. No hypothesis need however be made regarding the x's. They need not even be random variates, though in practice they are usually regarded as samples from some parent population.

In factor analysis, on the other hand, the basic assumption is that

$$x_i = \sum_{r=1}^{k} l_{ir} f_r + e_i \ (i = 1, 2, \ldots, p), \qquad (1)$$

where f_r is the rth factor (or common factor) and e_i is a residual representing sources of variation affecting only the variate x_i. The p random variates e_i are supposed to be independent of one another and also to be independent of the k variates f_r. The latter are assumed to be orthogonal and, without loss of generality, to have unit variance. All means may be taken to be zero. The variance of e_i may be denoted by σ_i^2. The quantities l_{ir} (and usually also the σ_i^2) are unknown parameters which have to be estimated. Attempts have been made to treat the individual values of the f's also as parameters, but this raises difficulties, a serious one being the fact that when the sample size tends to infinity so also does the number of parameters. We are here ignoring this approach, the practical point of which is not evident to us.

It is clear that the equations (1) are not capable of direct verification since the p variates x_i are expressed in terms of $p + k$ other variates which are not observable. The equations imply, however, a hypothesis, which can be tested, concerning the variances and covariances c_{ij} of the x's. These are in fact assumed to be capable of being expressed in the form

$$c_{ii} = \sum_{r=1}^{k} l_{ir}^2 + \sigma_i^2, \quad c_{ij} = \sum_{r=1}^{k} l_{ir} l_{jr} \ (i \neq j) \qquad (2)$$

Putting it in another way, we are setting up a hypothesis H_0 that the true covariance matrix can be expressed as the sum of a diagonal matrix with positive elements (either known or unknown) and a matrix of rank k with positive latent roots. The value of k is assigned in advance and must not be too large since otherwise H_0 might be trivially true. In the case where the σ's are known we must have $k < p$; but if the σ's are unknown and require to be estimated the condition becomes $p + k < (p - k)^2$.

Suppose now that we have a sample of sets of values of the x's and that we make use of the information provided by the sample variances and covariances. Then the likelihood function may be expressed in terms of these and of the population parameters, and maximization of the likelihood with respect to the unknown parameters will yield estimates of them. It may be noted that the estimates of the variances and covariances c_{ij} are not the usual sample estimates except in trivial cases where k fails to satisfy the inequalities given above. As an extreme example of such a case, suppose that $k = p$ and that the σ's are all zero. Then equations (2) can always be satisfied and H_0 thus makes no real-assertion about the variances and covariances; hence the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

maximum likelihood estimates of these are the usual sample estimates. The equations (1) are here merely equivalent to a component analysis.

For non-trivial cases a χ^2 test of the usual large sample type is available for testing the hypothesis H_0. Rejection of H_0 implies that the value of k is too small and that it must be replaced by some larger value.

As regards the estimation of the l's, no difficulty arises when only one common factor is assumed to be present. However when $k > 1$, the maximum likelihood method, though it provides a unique set of estimates for the c_{ij}, does not determine the l's uniquely. In a sense it determines the k-dimensional space in which the factors f_r lie; but it does not, and cannot by itself, determine where these factors lie. The position for the k factors is like that for the p components in a component analysis; there is an infinite set of possible orthogonal factors. In particular the maximum likelihood equations can be satisfied by an infinity of solutions all of which satisfy equations (2) and give the same value for the likelihood. It is clear that this must happen since in equations (1) the k variates f_r may be replaced by any other set obtained from them by an orthogonal transformation and this will produce a different set of l's.

In order to define the l's uniquely it is necessary to introduce $\frac{1}{2}k(k-1)$ arbitrary relations connecting them. These arbitrary relations may be chosen in such a way as to make the solution of the maximum likelihood equations as simple as possible. The effective number of unknown l-parameters is thus not pk but $\frac{1}{2}k(2p-k+1)$. The experimenter is of course under no obligation to accept this particular solution since it was selected merely on grounds of convenience. In fact he usually "rotates" or transforms to some other solution which he finds easier to interpret. Though to the statistician this may seem a very arbitrary procedure, in actual practice the values of the l's ultimately arrived at often throw light on the correlation structure of the variates and confirm *a priori* notions concerning the factors which are likely to be in operation.

One or two points have, for the sake of simplicity, been glossed over in the foregoing paragraphs. It may, for example, happen that the maximum likelihood equations have no solution in terms of real numbers, though this does not often occur in practice. Such occasional breakdowns of the maximum likelihood method are however known to occur in other fields of application and are not peculiar to factor analysis. Apart from cases of this sort the method has the usual large sample properties of maximum likelihood estimation and it can be shown that the estimates obtained are asymptotically normal.

To some people the multiplicity of possible values for the l's is rather disturbing; hence in conclusion it is perhaps worth considering briefly a possible method of avoiding the difficulty. In some investigations the experimenter may be able to visualise in advance the factors likely to be present and to be able to stipulate that certain variates should be independent of some of the factors. In other words it may be possible to assume *a priori* that certain of the l's are zero. Provided that there are a sufficient number of such zeros and that they are suitably placed this will cause the l's to be defined uniquely and the maximum likelihood equations to yield a unique set of estimates. The only drawback to this is that the equations become more complicated in form and far more difficult to solve, while the expressions for the standard errors of the estimated l's are in general very unpleasant.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

MEDICAL SECTION

In October, 1955, the Council approved a proposal that the sustained support for the Study Circle on Medical Statistics, which had since 1951 subsisted as a Group within the Study Section, though meeting separately, should be recognised by its constitution as a separate Medical Section. The regulations of this new Section are as follows:

Organization

1. The Section shall be constituted and carried on subject to the Bye-laws of the Society, and in accordance with these Regulations, made and subject to amendment by the Council of the Society. The Section shall be concerned with the application of statistical method to all branches of medicine.

Membership

2. Any Fellow of the Society and any other person who is approved by the Section Committee may become a member of the Section on making written application to the Section Secretary.

3. A Fellow of the Society shall not be required to make any payment for membership of the Section, but every other member shall pay an annual subscription of ten shillings, which shall be due in advance on January 1st, provided that any person who becomes a member and pays his initial annual subscription in the last quarter of any year shall not be required to pay any subscription in respect of the following year. Subscriptions shall be payable direct to the Honorary Treasurer of the Royal Statistical Society.

4. If any member of the Section has not paid his annual subscription by June 30th, the Secretary of the Royal Statistical Society shall advise him in writing of the default, and if the arrears are not paid up by September 30th, the defaulter shall cease to be a member.

5. Payment of an annual subscription in respect of membership of the Section shall not confer on the member any of the privileges of Fellowship of the Society.

6. Members who are not in arrears with their annual subscriptions may withdraw from membership by giving written notice to the Section Secretary.

The Section Committee

7. Subject to these regulations, all the affairs of the Section shall be directed by a Section Committee consisting of two Fellows nominated by the Council of the Society and six members elected by the Section.

8. (*a*) Two of the six members of the Section Committee elected by the Section shall not be eligible for re-election for the following session. The Committee shall decide on the members not eligible for re-election, and in doing so shall have regard to seniority on the Committee and degree of attendance. No elected member of the Committee may serve as such for more than four consecutive sessions.

(*b*) The Section Committee shall nominate members for election to the Committee for the following session, and these nominations shall be circulated to all members at least a month in advance of the date of the last Ordinary Meeting of the Section in the current session. Members of the Section may, if they so desire, nominate further candidates (being members of the Section) for election, in which case a ballot shall be taken at the meeting. Candidates so nominated must be proposed and seconded by members of the Section, and nominations must be received in writing by the Secretary of the Section at least a week before the date of the meeting.

(*c*) Casual vacancies in the Section Committee shall be filled by the Council and by the remaining members of the Committee respectively.

9. The Section Committee shall elect a Chairman and Secretary from among its members. These two officers shall retire at the end of each session but shall be eligible for re-election.

10. The Section Committee shall arrange meetings of the Section for reading papers, and holding lectures and discussions.

11. The Section Committee shall send to the Secretary of the Society not later than November 1st in each year (i) an itemized estimate of the total expenditure of the Section in the current session with a request for funds to meet such expenditure, and (ii) a detailed statement of the expenditure of the Section in the past session, with vouchers.

A report of the activities of the Section in the current session and some account of the papers, lectures, discussions and study groups arranged by the Section shall be sent to the Secretary of the Society not later than May 31st in each year.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REVIEWS OF STATISTICAL AND ECONOMIC BOOKS

CONTENTS

	PAGE
1.—*Braithwaite, R. B.* Theory of Games as a Tool for the Moral Philosopher	86
2.—*Bush, R. R., & Mosteller, F.* Stochastic Models for learning	87
3.—*Ostle, Bernard.* Statistics in Research	87
4.—*Haycocks, H. W. & Perks, W.* Mortality and other Investigations, Vol. 1	88
5.—*Wishart, J. & Sanders, H. G.* Principles and Practice of Field Experimentation, 2nd ed.	89
6.—*Kendall, M. G.* Rank Correlation Methods, 2nd ed.	90
7.—*Pompilj, G. & Napolitani, D.* Piano degli Esperimenti	90
8.—*Mannheim, H. & Wilkins, L. T.* Prediction Methods in Relation to Borstal Training	91
9.—*Sainsbury, P.* Suicide in London: An Ecological Study	92
10.—Political and Economic Planning. World Population and Resources	93
11.—*McCloskey, J. F. & Trefethen, F. N. (ed.)* Operations Research for Management	94
12.—*Cartter, A. M.* The Redistribution of Income in Postwar Britain	95
13.—*Galenson, W.* Labor Productivity in Soviet and American Industry	96
14.—*Dudding, B. P., & Jennett, W. J.* Control Chart Technique when Manufacturing to a Specification	97

1.—*Theory of Games as a Tool for the Moral Philosopher.* By R. B. Braithwaite. Cambridge University Press, 1955. 76 pp. 7¼″. 6*s*.

This little book is the inaugural lecture of the Knightbridge Professor of Moral Philosophy in the University of Cambridge (55 pp.) with the addition of some notes and an appendix. The greater part of the lecture with the whole of the appendix is devoted to a proposed solution of the non-zero-sum, two-person game in which the utility measurements of the two participants are not in related units. In the lecture the argument is carried out with reference to an example : Luke and Matthew inhabit adjacent rooms and have the same leisure hour, Luke plays classical music on the piano, Matthew improvizes jazz on the trumpet, and the sound insulation is negligible. How shall they arrange their playing to obtain in some sense an equitable apportionment of happiness? The solution is obtained by a delightful use of the arguments of projective geometry translated into game theory. The writing is excellent, the argument emerges with a lucidity that is all too rare in contemporary scientific writing. The appendix presents an algebraic treatment of the game-theory problem.

The major interest in the lecture resides in the fact that the utility measurements of Luke and Matthew are not in related units. This would happen in practice: for whilst it is in some degree reasonable to say the utility to Matthew of his playing the trumpet whilst Luke is silent is say, seven times his utility of being silent while Luke plays; it does not seem reasonable to say that Matthew's utility in the first case is ten times Luke's. Any solution that is offered must be invariant under a monotone transformation of either person's utilities. Nevertheless before a solution can be proposed we must have some means of assessing the relative advantage of one player over the other. Braithwaite does this by simply observing that a parabola has an axis and demonstrating some simple properties of treating this as a line along which Luke's advantage over Matthew is constant. The final solution to the problem involves collaboration between Luke and Matthew: in the numerical case given Luke will play and Matthew be silent on 17 evenings out of 43, on the remaining 26 evenings the roles will be reversed.

Is the solution satisfactory? First it is clear that no "solution" can possibly have the overwhelming appeal that von Neumann's minimax solution has in the zero-sum game: the most previous writers have reached is a set of reasonable solutions, i.e. nothing outside the set is a possible solution. With this proviso Braithwaite's solution has considerable appeal and appears to give intuitively sensible results. One possibly disquieting feature of the argument is that whilst it is obvious from the beginning that the players should co-operate (clearly Luke and Matthew would, if reasonable men, try to work out a joint plan) the solution involves arguing about what would happen if they did not. A possible criticism of the application to moral questions applies equally forcibly to the von Neumann theory in the simple case: namely the theory is a long-run theory, it does not deal with a single play. Of course Braithwaite's example is all right because the solution will apply for as long as Luke and Matthew live in adjacent rooms, but care is needed in applying it to questions which only arise once. Some military applications of zero-sum theory appear to the reviewer to be wide of the mark because they do not take account of the uniqueness of the play

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and the predilections of the opponent. In an infinite series of plays, of course, the players will have to "minimax", but in a single play a gamble may be worth while.

Braithwaite's book is a delight, and if there is a best solution to this problem his is most probably it. But is there?

D. V. LINDLEY.

2.—*Stochastic Models for Learning.* By Robert R. Bush and Frederick Mosteller. New York: John Wiley [London: Chapman and Hall], 1955. xvi, 365 pp. 9″, 72*s*.

It is natural that psychologists should try to find out how we learn, or that teachers should want to know how our learning could be improved. The authors of this book wish to contribute to just one aspect of this wide field. It is their view that learning is probabilistic, in that events change the probability of responses. One gathers that not all professional psychologists agree, but to a mathematical statistician the concept is of intriguing interest.

The authors "consider any systematic change in behaviour to be learning whether or not the change is adaptive, desirable for certain purposes, or in accordance with any other such criteria" (p. 3). Their book deals with a specific mathematical model. Its elements are a set of r alternatives A_j and a set of t events E_i. We start from a set of probabilities p_j of the alternatives. Now let a subject choose A_j. Then an event E_i occurs, and thereby the probabilities p_j are transformed into $p'_k = \Sigma u_{kj,i} p_j$ $(k = 1, \ldots, r)$. The process consisting of random choice, event, and transformation of probabilities is then repeated.

No stimulus-response theory of learning is assumed, and no definite position is taken on the issue of reinforcement theory v. contiguity theory (cf. section 9.2). However, Chapter 2 shows how the model could be deduced from considerations of stimulus sampling and conditioning.

Because the p'_k are again probabilities, and for other reasons connected with the intended applications of the model, the $u_{jk,i}$ are subject to certain restrictions. The relations between the alternatives chosen and the subsequent events are classified as follows : If the probabilities of the events are independent of the chosen alternative, the authors speak of "experimenter-controlled" events. When $r = t$ and each alternative induces a definite event, this is called "subject-controlled". (The theory is here related to Markov chains.) Finally "experimenter-subject-controlled" events are identified with an alternative and the occurrence of one out of a set of possible outcomes. This case involves probabilities of alternative-cum-outcome pairs.

Having constructed the model, one will be interested, *inter alia*, in the successive distributions of probabilities and in their asymptotic values, These are problems dealt with in the eight chapters of Part I : "The Mathematical System and the General Model". Some interesting theorems specifiying intervals within which the values of the p_j are ultimately trapped are mentioned without proof because these "are long and require advanced methods" (p. 98), which would exceed the scope of the book. Chapter 6 : "Approximate Methods" introduces a Monte Carlo technique by considering a hypothetical animal called "stat-rat" (p. 129) which runs repeatedly through a sequence of trials. Chapter 7 treats "perfect" learning, where the limits of the probabilities are zero and one, and Chapter 8 yet another special case, where the matrices $T_i = (u_{kj,i})$ are commutative.

Part II: "Applications" considers, in 7 chapters, the problems of identifying (i.e. interpreting) the parameters, of estimating their values, and of testing their goodness of fit. In this connection various experimental observations are described which are, according to taste, either illustrative examples or indications of the ultimate aim of the study.

There are references after each chapter and a glossary of symbols. The reviewer was greatly helped by the excellent summaries to all chapters and he recommends them to the reader. But he is unable to recommend a new word in the first sentence on p. 201, which reads: "The solution . . . is $p = 0, 7$, as we have intuited!"

S. VAJDA.

3.—*Statistics in Research.* By Bernard Ostle. Ames : Iowa State College Press, 1954. xiv, 487 pp. 9″. $6.95.

The sub-title of this book is "Basic concepts and Techniques for Research Workers", and the subject-matter has been taught for some years at Iowa and Montana State Colleges. It is a very long book, of fifteen chapters and fifteen appendices. Broadly, after brief introductions on the role of statistics in research and the basic statistical concepts, the book attempts to deal with the whole apparatus of modern statistical methodology, random sampling distributions for measurement data, χ^2 and other methods for attribute data, tests for means and variances, regression and correlation analysis, analysis of variance and covariance, ending with a brief chapter on other statistical techniques and one on the design of experimental investigations.

The book, therefore, does not break any new ground, but is content to follow the familiar pattern of books on statistical methodology which, because they prove no mathematical theorems,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

are essentially didactic in tone. There is mathematics in the book in the sense that the author quotes the frequency functions for the distribution laws of t, χ^2, F, etc., but these are merely stated. The book is hardly likely, in consequence, to appeal to the mathematical statistician. Its success will depend upon the extent to which the non-mathematical reader will be able to follow the logic of the author, and to accept the necessity, and the detail, of all the tests recommended. The opinion of the reviewer is that the author has not wooed his public sufficiently at the outset. His style of writing in Chapter 1 is not very convincing, while by page 28 he is throwing at the reader the mathematical formulae already referred to, with little or no explanation. A number of very difficult concepts are introduced too early, such as tests of hypotheses, Types I and II errors and confidence intervals. The distributions of a mean, of the sum or difference of two means, of the sample variance, of t, of F and of the range, are all given experimentally by choosing 600 samples of 5 from a normal population with mean 4,000 and standard deviation 1,000. From the unorthodox presentation of the results in the form of frequency distributions with unequal class intervals, the reader is expected to see the fundamental distribution results in operation, and to accept vague statements such as "the differences between our sample means . . . seem to vary about twice as much as did the single means." The mathematically inclined reader is asked to check that $\sigma^2_{\bar{Y}} = \sigma^2_Y/n$; others are not told how to convince themselves from the frequency data presented.

If the reader is able to break into the main sections of the book, he will find that, for the most part, the methodology is correct and sufficient, and that there is not much left out which is necessary to an introductory course of instruction in the subject. Examples are given everywhere, even if some of them seem to lack an air of reality, and each chapter ends with a number of set problems. The treatment is sometimes over-lengthy; for example the chapters on analysis of variance occupy 150 pages, of which 11 lines are given to Incomplete Block Designs and 7 to Lattice Designs.

The book has a number of irritating sections and naïve phrases, and can be criticized in places. A few examples will indicate what is meant. The binomial distribution is introduced in a very unorthodox and difficult manner. Two approximations are used on pp. 57 and 58 without it being stated that they are the same, provided, that is, the continuity correction (which the author seems to have overlooked elsewhere) is applied to χ^2. On p. 74 the reader is told to double the normal probability from Appendix 4 for use with the $\sqrt{(2\chi^2)}$ approximation, but surely this is incorrect. Many readers will be baffled by figures which are stated to come from the tables in the Appendix, but which cannot be found therein. They are sometimes given in the text to more places of decimals than in the tables themselves. The fact is that no mention seems to be made of the necessity of interpolation, nor of the methods to be adopted. Finally, two examples may be given of things which might have been said differently. Referring to the variance-ratio distribution, "The reader should not gain the impression that unity is a 'holy' number when values of F are being discussed". The author discusses covariance in a pig-feeding trial on pp. 390–392, finds that original and adjusted treatment means are not significantly different, and only at the very end shows that the error regression is not significant either. He goes on: "In such a case it does not appear that we are justified in applying the covariance technique—perhaps that is why it produced a conclusion no different than that obtained from a simple analysis of variance of the gains in weight".

J. WISHART.

4.—*Mortality and other Investigations*, Vol. I. By H. W. Haycocks and W. Perks. Cambridge University Press (for the Institute of Actuaries and the Faculty of Actuaries), 1955. ix, 164 pp. 8½″. 20*s*.

This is the first volume of a further book in the series, intended principally for actuarial students, that is being published for the Institute of Actuaries and the Faculty of Actuaries. Its content is thus largely governed by the examination syllabuses of those bodies and it would not be fair to criticize the authors for the selection of the topics included, nor can a final judgment be passed on this until the contents of the second volume are known. The present volume takes the study of the derivation of rates of mortality from the most elementary starting point up to the use of the simpler forms of the so-called "continuous" exposed to risk formulae. It then discusses sickness rates and, very cursorily, other rates commonly used by actuaries. There follows a chapter on graphic graduation (other methods of graduation being left for Volume II) and graduation tests. The last two chapters deal with the methods developed by George King for deriving life-tables from population data and with the construction of the series of English Life Tables.

The book opens with an introductory chapter on "actuarial statistics". This, though it may not mean a great deal at first reading to the beginner in the subject, will be of considerable interest to the statistician who is not an actuary, for it will give him a clear insight into the attitude of mind which the practising actuary brings to bear on statistical theory.

The most important part of the book is contained in Chapters 2 to 5. In these the subject of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

"exposed to risk" is approached in an unusual but most stimulating manner. The authors begin by considering how to find the rate of mortality for a group of lives observed for a period of one year; the problem is brought to its simplest terms by assuming that there are no entrants or leavers. This is followed by the longest and most vital chapter of the book in which it is shown how these group rates of mortality may be analysed by age. The authors consider in turn the life-year, the calendar-year, the policy-year and the census type of investigation, and end the chapter with some most helpful diagrams. The next chapter considers the combination of two or more years of experience and here the value of the authors' approach shows itself, for it now becomes clear that the traditional methods of obtaining the exposed to risk, i.e. the "continuous" or, as the authors call them, "difference" formulae, are no more than means of saving labour when the experience covers more than one year. Precisely the same results are obtained from the continuous formulae as would be obtained if each year of such an experience were treated separately and the resulting exposed to risk for each year added together, and it is a great virtue of the book that this fact is made abundantly clear. To conclude this section Chapter 5 deals with entrants and leavers, and the reviewer would agree with the authors that "these complications are little more than practical nuisances" and are rightly left to the end.

Though the general line of argument adopted in these four chapters has many advantages, the reviewer considers that the student coming to the subject for the first time will find them somewhat indigestible. What appears to be needed is some leavening of the theory with practical examples. It would have been most helpful if particulars such as date of birth, date of entry and date of death had been given for a score or so of imaginary lines and the authors had shown how these cases would be treated in each of the various types of investigation they consider. Students of this essentially practical subject seem to learn much more from one well-chosen example than from the most carefully considered theoretical exposition. It is also to be regretted that there is no word about the actual collection of the data, a matter that in practice abounds with difficulties and pitfalls —but perhaps the authors regarded it as outside the scope of this volume.

The rest of the book may be considered more briefly. Chapter 6, which is mainly about rates of sickness, contains an interesting discussion of the various measures of sickness; the terminology used for these measures by various writers in the past has not been noticeable for consistency, and this chapter may serve as a reminder of the necessity for careful definitions in this field. The following chapter on graduation runs into the usual difficulty that a graphic graduation is easier to do than to describe, whilst the section on tests of a graduation contents itself with describing the tests and says little about the interpretation of the results; it might have been emphasized that this interpretation must be determined largely by the purpose for which the graduated rates are to be used. The last two chapters are a convenient summary of methods commonly used in obtaining mortality rates and life-tables from population statistics.

This volume may be read with advantage by any statistician, or for that matter any actuary, who is interested in the subject of mortality rates. The approach adopted will almost certainly give the reader a fresh insight into the fascinating subject of exposed to risk, for the authors have stripped away many of the complications of the subject and laid bare the essentials. Whether or not this volume is an ideal one for the beginner is perhaps more doubtful—he may find it difficult to apply the theory to practical examples without more help than the authors give him—but there is no doubt that, if he masters the authors' exposition, he will have in his grasp all the basic principles he needs for his further studies.

LESLIE V. MARTIN.

5.—*Principles and Practice of Field Experimentation.* By J. Wishart and H. G. Sanders 2nd ed. Cambridge: Commonwealth Bureau of Plant Breeding Genetics, 1955. vii, 133 pp. 8½". 21*s*.

This book provides one of the best introductions to experimentation for the practical worker in the field of agriculture. In 80 pages, Dr. Wishart covers the basic statistical techniques, describes randomized blocks and Latin squares and gives an outline of the more advanced techniques in current use—factorial experiments, split plots and incomplete block designs. Sampling techniques and the analysis of covariance are also treated in general terms. There is no mention of experiments extending over more than one season, but within its chosen limits the only serious criticism that can be raised is the somewhat undue emphasis placed upon significance testing, a lack of balance that is shared by almost all the current elementary textbooks.

Dr. Sanders's contribution, though it contains little of direct statistical interest, is probably the more valuable part of the book since it presents information that is extremely hard to come by elsewhere. The practical flavour of this section is very marked and the author is clearly writing from a considerable (and sometimes bitter) experience. It is perhaps on a basis of this experience

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that Dr. Sanders asserts with some emphasis in his opening chapter that statistics and formally designed experiments must be kept in their place and should play little part in the preliminary stage of an investigation; at least one statistician has found that a reversed emphasis on the drawbacks of "demonstrations" and "observation plots" is often in order. M. J. R. HEALY.

6.—*Rank Correlation Methods.* By M. G. Kendall. 2nd ed. London: Charles Griffin, 1955. viii, 195 pp. 9″. 36*s*.

This is the second edition of a book which although published only in 1948 has established itself in statistical literature. The first edition, written by one who has contributed more than anyone else to this particular branch of statistics, was welcomed as filling a long-felt want. The second edition consists of a revision of the first with some new material interpolated; the original plan of the text is left unchanged. The principal changes appear to be the insertion of some algebraic matter in Chapter 2, including Daniel's inequality and the inequality of Durbin and Stuart, the application of rank correlation methods to ordered contingency tables (Chapter 3), the introduction of Hoeffding's inequality into the text (Chapter 4), a brief discussion of the distribution of Spearman's ρ in the non-null case (Chapter 5), incomplete rankings (Chapters 6 and 7), and the relation between ranks and variate values (Chapters 9 and 10). A new chapter on "Some Further Applications" has been inserted at the end. A table of the random permutations of the first twenty natural numbers has been added in the Appendix.

The changes are not of sufficient magnitude to require those statisticians who possess a copy of the first edition also to buy the second. Some of them are no doubt ephemeral and will be altered in their turn in the third edition. Either in its first edition or in its present form the book is still indispensable to those statisticians interested in rank correlation methods

F. N. DAVID.

7.—*Piano degli Esperimenti* . . . By Giuseppe Pompilj and Diego Napolitani. Rome: Supplemento a "La Ricerca Scientifica", Anno 24°, 1954. 206 pp. 9½″. *L*2000.

This publication comes from the National Institute of Nutrition in Italy. It deals with the design of experiments and the statistical analysis of the results, with special reference to biological experimentation. There are three sections, each consisting of one long chapter. The first begins with an exposition of the laws of probability, in their application to experimentation. The grouped frequency distribution is then described, by means of an example, and the mean and variance defined, although without making any distinction, at any rate in notation, between population parameters and sample estimates. The use of the term "deviance" for the sum of squares of deviations from the mean is unusual (and equally the use of "codeviance" for the sum of products in correlation analysis). The chapter ends with the detailed calculation of mean and standard deviation, and the fitting, not very successfully because of the nature of the data, of a normal curve to the authors' example.

The second section is taken up with descriptions, illustrated by examples, of the main statistical tests—(1) Pearson's χ^2 applied to test the goodness of fit of a normal curve to grouped data, and in contingency tables, (2) "Student's" t-test illustrated in the usual way, but with an extension to orthogonal comparisons in more than two samples, and (3) Fisher's F-test explained for the cases of gradually increasing complexity up to those of unequal numbers in the sub-classes. The analysis of covariance is described at this stage, but, so far as can be judged, the particular way in which the F-test is used here is not correct. In this chapter the F-ratio is defined only for the usual analysis of variance situation, for which the "one-tail" probabilities are appropriate. It is surprising, therefore, to find that the table of F at the end gives what are usually described as the 2·5, 0·5 and 0·05 per cent. points instead of the 5, 1 and 0·1 points. Reference is made to *Statistical Tables* by Fisher and Yates, but the authors' tables are not to be found therein. There is a use for the "two-tail" probabilities, but not in the situations dealt with by the authors.

The third section takes us through the usual sequence of experimental designs, and the corresponding statistical analyses, from the completely randomized design to cases of confounding and incomplete blocks. There are plenty of worked examples, most of them taken from other published sources; very few illustrations are from Italian sources. To sum up, the book will be useful in the sense of providing in the Italian language a summary of the main work done during the present century in this important subject, but it is not as authoritative as certain well-known English and American texts, and should be examined critically by the reader with a view to amendment at certain points. In addition to the F-table already referred to, there are tables of the normal integral, the χ^2 and t distributions, and of squares, square roots and reciprocals. A summary of contents is provided in four languages.

J. WISHART.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

8.—*Prediction Methods in Relation to Borstal Training.* By Hermann Mannheim and Leslie T. Wilkins. London : H.M. Stationery Office, 1955. viii, 276 pp. 9½″. 17*s.* 6*d.* net.

The main object of the research described in this report was to determine whether it would be practicable to construct " prediction tables " of the type now widely used in the United States for estimating the likelihood of success or failure in the case of youthful offenders sentenced to a training institution. A review of previous inquiries (mainly American) reveals two different modes of approach : most investigators have assumed that the best predictions are to be obtained by including as wide a variety of prognostic factors as possible ; but the more recent prefer to rely on a small number of factors, assigning to each its own differential weight. In either case, as Dr. Mannheim and Mr. Wilkins point out, the statistical procedure adopted has, as a rule, been extremely crude ; and no allowance has been made for the inevitable overlap between different factors.

In this country a similar problem continually confronts both educational and industrial psychologists, who have to decide what type of school or vocational training will be most suitable for a particular candidate. The procedure commonly used for such purposes has been an adaptation of Pearson's method of multiple correlation—most frequently some form of factor analysis or discriminant function. Although they do not explicitly refer to British experiments along these lines, the methods proposed by Dr. Mannheim and Mr. Wilkins are of the same general nature. But they have incorporated several interesting modifications of their own to meet certain difficulties which so often perplex the investigator in practical studies of this kind.

A sample of just over 700 Borstal youths was available ; and the basis for prediction was "the information matrix consisting of the data available at committal and the data that became available during training ". Some of the information was expressed by continuous variables; most of it was in the form of dichotomous classifications. According to the type of items to be correlated, therefore, the investigators used an ordinary product-moment coefficient, a biserial coefficient, or a coefficient of mean contingency corrected for " restricted limits". The calculation of the "weights" was divided into two stages. "We first considered all attribute data on its own, ignoring variables, and the attributes were weighed to provide the 'best' estimator of the criterion. This equation was then treated as a 'score', and put into the variables matrix as a further variable. The result of the final solution was applied to the 'score' and the variables, in order to derive the weights; and approximations were then made to the nearest integer".

From the composite assessment thus computed it is possible to estimate the probability of success or failure for any youth about whom the usual information is available; and a graph is provided for this purpose. To simplify the results still further, the investigators suggest that the final prediction should be expressed in terms of five 'risk categories'; for these the probability of failure is approximately that shown by the figures in brackets: A (1 in 8), B (1 in 3), X (1 in 2), C (2 in 3), D (7 in 8). But here another difficulty arises in investigations of this type. The " central category of unpredictables (X)"—i.e. "those for whom the likelihood of success and of failure is almost equally balanced"—proves usually to be the largest of all. Accordingly, for these more doubtful cases the investigators have constructed a second multiple regression equation: "the factors which failed to appear in the first equation, but for which the zero order coefficient was significant, are now put back again for the central group".

On comparing the "predictions" thus derived with the actual record of success or failure for the youths in the same sample, the investigators obtain a correlation of 0·50, when behaviour during detention is included in the sample, and 0·46, when it is not. Since most of the American writers make no use of differential weights, they have also calculated figures for prediction derived by equal weighting; the correlations so obtained are said to be "slightly less", though actual coefficients are not stated. On applying the same weighting equations to information recorded a year or two later for an entirely fresh sample, they obtain a correlation which is virtually the same as before namely, 0·48.

The report includes much additional material of great interest to the criminologist, and concludes with a suggestive chapter on the future of criminological research. At the present stage, however, the main purpose of the investigators has been to develop a statistical technique which should be both workable in practice and sound in theory. Considering the innumerable difficulties that inevitably beset all such studies, the whole investigation has been excellently planned and skilfully carried out. Indeed, as the writers themselves have indicated, the procedure which they have evolved might well be applied, with perhaps a few minor modifications, to a great variety of administrative problems in spheres outside that of criminology. The presentation is clear, simple and logical, and marked from beginning to end by wise common sense. The work itself fully justifies the recent decision of the Home Office to embark on a series of such researches, and sets a high standard for the volumes that are to follow.

CYRIL BURT.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

9.—*Suicide in London: An Ecological Study*. By Peter Sainsbury. London: Chapman and Hall (for the Institute of Psychiatry), 1955. (Maudsley Monographs No. 1.) 116 pp. $9\frac{3}{4}''$. 15*s*.

In no sphere of medicine is the interaction between the individual and the community of greater importance than in psychiatry, but it is difficult to find, in this country at all events, factual evidence bearing on the subject. Dr. Sainsbury, in his study of the ecology of suicides in London, has made a notable contribution to this important, albeit neglected, field. It is all the more noteworthy because he is well aware of the difficulties in investigating socio-economic aspects of a large conurbation such as London with inadequate population statistics.

The investigation was: "to test the hypothesis that where social mobility and social isolation are pronounced, community life will be unstable, without order or purpose, and this will be reflected to a greater or less degree in the suicide rates" (p. 30). It was conducted in two parts: (1) by a "statistical correlation of suicide rates in the twenty-eight Metropolitan Boroughs and the City with selected indices of their social characteristics"; (2) by an analysis of the case histories of 409 suicides investigated by the Coroner of North London over a 3-year period (1936–38).

The results are presented in the form of tables and maps. An Appendix includes the basic data tables on which the work depends. The list of references is valuable. The shortcomings of ecological studies are discussed in the introductory chapter, which includes an admirable critical review of the literature on this subject. In trying to correlate socio-economic data relating to Metropolitan Boroughs with the suicide rates of those Boroughs the investigator is not a free agent since he is dependent on diverse sources for many of the figures on which he bases his conclusions. This is particularly true of one feature which Dr. Sainsbury investigated, namely, *social mobility*. (It should be noted that Sainsbury uses the term to denote transiency of the population in contradistinction to Glass *et al.* (1954) who use it to denote movement up and down the social hierachy). The two indices, expressed as a proportion of the total population in the borough, used by Sainsbury to measure social mobility were : (a) the number of "immigrant" persons (i.e. not London born); (b) the sum of the total population coming into and leaving the borough daily. These are at the best crude measures, particularly as it is not possible to compare the age, sex and marital status distributions of the mobile and static populations. The first index may be regarded as a reasonably reliable measure of the exotic population in the borough, but the second has little to commend it.

The daily exit from a borough can possibly be used to measure the mobility of the Borough population, but the number of persons coming into the Borough does not affect the mobility of its resident population. Boroughs with a large traffic are those such as Kensington, St. Marylebone and Westminster, the main shopping areas, Holborn and St. Pancras wherein are located the University of London, various colleges, schools for further education, the main commercial and business offices, main-line railway stations.

Dr. Sainsbury has shown that there are differences between the Metropolitan Boroughs both in respect of suicide rates and of socio-economic features, that there is statistical correlation between high suicide rates and Boroughs with favourable socio-economic features and that the converse also holds good. The acceptance of statistical correlation as evidence of causal relationship is dangerous. It is possible, for example, to use Dr. Stocks's cancer figures (1947) and find that cancer mortality for all sites except stomach show a similar pattern to the suicide rates. The "cancer" maps of the County of London produced by Stocks are, in fact, very like the "suicide" maps drawn by Sainsbury. Indeed, when the Boroughs are ranked in descending order of suicide rates all but one of the ten in the top quartile have a cancer mortality which is above the average for London and all of the seven Boroughs in the bottom quartile have a cancer mortality below the average for London. It is true that Dr. Sainsbury found that *alleged* malignant disease was stated to be a contributory factor in 14 out of 390 suicides, yet this would hardly be of sufficient importance to account for the apparent association between malignant disease and suicide. Even if the correlations found by Sainsbury are regarded as indicating causal relationships, and this is a point on which there is bound to be conflicting opinions, it is hardly possible to concur, without further evidence, with the statement (p. 89) that "if a community is mobile human relationships are perfunctory, and those values and standards are lacking by which it controls its members, and by which they order their lives".

If this work is regarded as an initial survey of a complex problem, and in his discussion the author intimates that this is his own view, then Dr. Sainsbury has performed a very useful task. While it is necessary to criticize details of the technique employed it is very desirable to praise the general approach and treatment of the subject. Medico-statistical surveys of this type are in their infancy. Each research worker is a pioneer and must use the crude tools at his disposal. Indeed, without such pilot studies it will not be possible to fashion more sophisticated techniques which will appeal to the professional statistician.

Some interesting results emerge which will certainly stimulate further research into this subject. Poverty as such is not frequently a cause of suicide, but a sudden change from affluence to compara-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tive poverty is. Overcrowding showed no relationship to suicide rates. Loneliness was frequently recorded as a cause for suicide. One quarter of the suicides were said to have had some abnormality of personality. Physical illness contributed to suicide in one-quarter of the cases, and among the suicides tuberculosis and cancer were more common than among the general population. Mental illness was the principal cause in one-third of the suicides, and this suggests that ecological studies of mental disorders will throw more light on the suicide problem and *vice versa*.

In all, Dr. Sainsbury is to be congratulated for a careful and thought-provoking investigation of an intricate subject, which has been recorded in an easy style.

V. NORRIS.

References

GLASS, D. V. (1954) Ed. *Social Mobility in Britain*, London, Routledge & Paul.
STOCKS, P. (1947) *Regional and Local Differences in Cancer Death Rates*, London, H.M.S.O.

10.—*World Population and Resources.* A Report by P.E.P. London. P.E.P. (Political and Economic Planning), 1955. xxxvii, 339 pp. 8½″. 30*s*.

The Malthusian problem in its modern, world-wide form has not suffered from lack of attention in the Western World in recent years. Many—probably too many—works on its various aspects have been published, often expressing strong views on matters that have not so far been capable of proper measurement or analysis. The idea that there should be a more scientific investigation of the interrelationships between population and resources is a comparatively new one, to which attention has lately been drawn, under the auspices of the United Nations Organization, by the issue of a volume entitled *The Determinants and Consequences of Population Trends* and by the private circulation to interested persons of a list of fifty demographic and economic questions deserving of serious research. Until the fruits of this research are ripe, what is most needed is a more practical and responsible discussion of the available data concerning the growth of man's numbers in relation to the resources that have been, and can be, developed.

The Report now issued by *Political and Economic Planning*, which is the work of a Group of natural and social scientists (whose names are not given but who are undoubtedly prominent in their own fields) should provide a firm foundation for both action and discussion. In it, the difficulties of the present time and of the years soon to come are outlined in such a manner that particular emphasis is laid on the problems of the underdeveloped areas. It is in these areas that population presses the most heavily on resources and that the advance towards higher living standards is the most difficult to achieve. It is in respect of these areas also that the least is known about the interrelationships between human numbers and the supply of the necessaries of life. The tendency in most writings has been to treat the two functions too independently; although the P.E.P. Group have sometimes been obliged to adopt a similar form of approach they have shown themselves fully conscious of the need for fresh thought, and have devoted an entire chapter to a consideration of the problems of interrelationships.

The first half of the Report examines the evidence, first for the world as a whole and then for a representative selection of countries and islands. The authors conclude that population problems exist everywhere and at all times; in other words, every human community has perpetually to be watching its housekeeping budget. These problems are, however, variable in intensity and even in their character, and the possible remedies are so difficult and controversial that there is much temptation to regard them as insoluble; nevertheless, as the Group say, they are far too important to be neglected. The second half of the Report describes and comments on the population policies of those few Governments that have progressed so far as to have a positive policy on the subject. The concluding section gives details of the programme of action that it is considered should provide the necessary background for the formulation of policies to improve man's circumstances, especially in the underdeveloped areas. Each part of the Report has been made as far as possible self-contained, and this arrangement, together with the lucidity of the writing, is most effective in conveying the Group's point of view clearly to the reader, even if repetition is not entirely avoided.

Some writers tend to argue away the troubles of today and tomorrow by looking ahead to a remote time when scientific advances will have harnessed all the materials of the earth in the most effective manner for man's benefit. Whether or not such a time will come, speculation of this kind is of little help in ameliorating the lot of the human race either today or during the next two or three decades. The prospects for the more immediate future are the rightful concern of the P.E.P. Group, who after a short survey of current world population trends, give an account of hoped-for improvements in agriculture, in the winning of mineral resources and in the provision of industrial energy. Discussion of these matters in world terms has many limitations, and the text

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

therefore goes on to give nineteen short surveys of important or typical areas among which certain islands afford the most striking examples of Malthusian dangers. It might have been preferable to have provided fuller accounts and more thorough analyses of a smaller number of countries, but although occasionally over-simplified the presentation is pithy and on the whole very effective.

In a section entitled "Nation and Family" the actions of Governments that place one or other of these ideals high in their population policies are contrasted. Among those policies which, perhaps, have regard more to the nation than to the family there are two designed—in their different ways—to promote births (those in France and Russia) and two designed to limit births (those in India and Japan). Sweden and the United Kingdom are cited as examples of countries where the family is the major object of concern. The formulation of population policies is a task to which many more governments will have to address themselves in the coming years, and in making this latest survey of tendencies the Group have therefore performed a useful function.

State intervention can, if made drastically enough or at the appropriate moment, have a great influence on population, even if it runs contrary to established belief and custom. Sometimes voluntary action can also be powerful. Having shown these facts to be true, the authors of the Report have gone on, in their concluding remarks, with some confidence to outline the steps that will be required to keep the Malthusian devil at bay in the coming decades. First of all, the great unawareness of population problems that still exists in the world must be abolished and supplanted by the spread of enlightenment. Secondly, research into the interrelationships between the growth of numbers and the growth of resources, and into the regulation of human fertility, must be vigorously pursued. Finally, surveys should be made of all areas where population is outstripping food and other supplies, and in these areas technical aid and other programmes of assistance should be adapted to the double task of promoting the development of resources while at the same time restraining population from making a corresponding or greater increase.

The conclusions and recommendations of the Group seem fairly to sum up and round off the book. While the forecasts upon which much of the argument rests may be fallible, and while the authors may not be entirely free from bias—both of these possibilities are freely admitted in the Report—it can at least be said that a valiant effort has been made to treat an awkward subject with the logic and impartiality that its great importance demands.

P. R. Cox.

11.—*Operations Research for Management.* Edited by J. F. McCloskey and F. N. Trefethen. Baltimore: Johns Hopkins Press, 1954. London: Cumberlege. xxiv, 409 pp. 9″. 60*s*.

The term "Operational Research" was coined before the second world war to distinguish those radar research scientists who found they could advise the armed services on the operational use of radar and its integration within the general scheme of the defence of this country. This does not mean that operational research as such had not been carried out earlier, but this was when it began to be recognized as a distinct entity. The war gave a great impetus to the development of operational research methods in the military sphere, and with the end of the war the recognition of the value of operational research spread to industry. These remarks apply to the United States as well as to this country, and in each there is a flourishing society devoted to collating operational research and to providing a forum of discussion for its practitioners. In the United States a number of universities, research institutes and institutes of technology have operational research groups doing advisory work for industry. Many of these institutions run regular training courses in the subject as part of their curriculum and it is now possible to obtain a doctorate in operational research at the Case and Massachusetts Institutes of Technology. In this country regular courses have been held at Birmingham University, but as yet academic institutions have not interested themselves in operational research on anything like the scale of the Americans.

Some years ago Morse and Kimball wrote their now celebrated book on operational research.* This covered mainly the military applications of the subject. What has been lacking until now has been a text-book of operational research from which the intelligent layman can see how far its techniques and applications warrant it being established as a distinct entity, to which the practitioner may turn for technical advice and information. Such a book now appears from Johns Hopkins University and is the subject of this review.

For some years Johns Hopkins have run a regular weekly seminar in operational research to which visiting lecturers have been invited to give papers. These papers have been published by the University as monographs and some have appeared in the *Journal of the Operations Research Society of America.* This new book brings together in one volume a number of these papers. The book inevitably gains advantages and suffers drawbacks. It is clearly an advantage to have

* Morse, P. M. & Kimball, G. E. (1951), *Methods of Operations Research.* New York: Technology Press of Mass. Inst. of Technology and John Wiley. London: Chapman & Hall.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the various sections of a pioneering book in a new field written by people who can speak with authority on the particular subjects they cover. The disadvantage is that inevitably the linking between successive chapters is not well established and the common theme of operational research methods is not defined with sufficient emphasis for the layman to grasp at first sight.

The book is divided into three main sections. The first is a general one with an extremely interesting account of the history of operational research from 1936 onwards. It is perhaps a pity that earlier historical examples of operational research are not given. Other chapters in this section cover operational research as a profession, and as a science, as well as its organization and relationship to management consulting.

The second main section is devoted to the methodology of the subject. This is covered by chapters on statistics, queueing theory, information theory, suboptimization, symbolic logic, computing machines, linear programming and game theory. These chapters are particularly well written and give concise accounts which can be used as introductions to the particular techniques.

The final section gives a number of case histories. One, a study of operational research in agriculture by C. W. Thornthwaite, is an admirable example of how a simple technical job, in this case to advise on supplemental irrigation, can become subsidiary to a much broader and more important problem. One omission in this and the preceding section is any discussion of communication problems.

The dividends accruing from a "mixed team" approach in operational research are by now well known but the danger of specialist scientists trying to cast an operational research problem in the image of their own specialization must be avoided. It is, therefore, necessary in a text-book of this sort to establish the primacy of the type of problem rather than the type of solution employed. In this book we have two distinct sections, one on technique and one on problems in different industries. The general theme of operational research might have been established more clearly had the second and third sections been reversed in order. The second section would then be devoted to problems of operational research, and these could well be better classified according to their general type than to the industry in which they occur. Such a classification by type is given in this book by Goodeve in his chapter on operational research as a science. He classifies problems in three main categories as follows:

I. *To do with people alone.*—(*a*) Organization and management. (*b*) Absenteeism and labor relations. (*c*) Economics. (*d*) Decisions of individuals. (*e*) Market research.

II. *To do with people and machines.*—(*a*) Efficiency and productivity. (*b*) Organization of flow in factories. (*c*) Methods of quality control, inspection and sampling. (*d*) Accidents. (*e*) Organization of technological change.

III. *To do with movement.*—(*a*) Transport. (*b*) Stocking, distribution and handling. (*c*) Communications.

The sections on techniques and methodology would then come as a final section which would give an introduction to the methods being used in present day operational research.

The authors are, however, to be congratulated in bringing together in one volume so much that is new and stimulating both to the operational research worker and the scientist in contiguous fields.

B. H. P. Rivett.

12.—*The Redistribution of Income in Postwar Britain: A Study in the Effects of the Central Government Fiscal Program in* 1948–49. By Allan M. Cartter. Yale University Press (London, Geoffrey Cumberlege), 1955. xi, 242 pp. 9¼″. 40*s*.

This is a book of scope and objective which make it of direct interest to a wider audience than can usually be expected for a purely statistical study. The author has attempted to measure the redistribution of incomes in post-war Britain with particular reference to the effect of the legislative programme for the year 1948–49.

He has divided his study into two parts. The first is mainly concerned with the final results and includes an introductory chapter on the theoretical problems and a concise summary of the general conclusions. This Part is well designed for the general reader. Professor Cartter has set out the results of his statistical calculations in clear tables and text, has analysed these data to seek answers to a series of very pertinent questions and has interpreted the answers with a forthright conviction.

The second part is addressed to the technical reader and describes the assumptions, raw material and statistical methods underlying the tables in Part I. It is to be hoped that the general reader will take the trouble to read enough of Part II to realize how much personal ingenuity has been applied

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to derive the statistics of the earlier chapters, and to respect Professor Cartter's warnings about the statistical accuracy of his results. The technical reader, however, may feel that he has not been so well served as the general reader, for the author is not invariably explicit about the details of his estimates. Although it is made clear enough in general terms what kind of material he has used and to what kind of processes he has submitted it, the reader who would like to repeat some of the calculations, and possibly to check some of the intermediate results and implications by alternative methods, will find a number of blanks in the narrative. It is nevertheless true that Part II is well documented and is even richer in statistical data than Part I. It illustrates with force and candour the considerable degree of guesswork involved in these estimates of income redistribution.

The author concludes that about 13 per cent. of the national income of Great Britain was redistributed in 1948–49 and that this compares with a percentage of rather less than 9 per cent. for 1937 and under 8 per cent. for the United States in 1946–7. He measures inequality by an index which reflects the percentage of incomes which would have to be shifted from high to low incomes in order to achieve complete equality of distribution. He estimates that this percentage was over 40 per cent. at the beginning of the century and 24 per cent. in 1948–49: and he judges that a marked decline in the percentage for disposable incomes between 1937 and 1948 was entirely due to the redistributive effects of direct taxes and transfers. " Britain went about one-third of the way to establishing perfect equality in 1937, but almost half-way in 1948–49, while the United States went only one-sixth and about one-quarter of the way in the respective pre-war and post-war years " (p. 86). He considers also the distribution of wealth, and concludes from death duty returns that in 1936–38 nearly 77 per cent. and in 1948–49 over 75 per cent. of private capital would have had to be redistributed in order to equalize capital holdings among the adult members of the population.

This is a thoroughly readable report on a competent and worth-while piece of statistical research.

P. M. DEANE.

13.—*Labor Productivity in Soviet and American Industry.* By Walter Galenson. New York : Columbia University Press, 1955, London: Cumberlege. xiv, 273 pp. 9¼″. 45*s*.

This book was prepared as part of the large programme of studies by the Rand Corporation for the United States Air Force. The author, therefore, it may be assumed, was not handicapped by any lack of money, staff, or documentation. We are told that a good deal of information has been extracted from Russian technical journals which had not been studied before by outside researchers.

Considering these circumstances, one is bound to notice that, while the book contains some interesting information, there are significant omissions. In 1955 the Rand Corporation has published a book of studies on productivity which come to an end in 1938. Studies of the economic history of Soviet Russia in the 1930's are interesting, but the title of the book should not be allowed to convey the impression that it deals with present-day Russia. The author makes a half-hearted attempt to include 1950 in the comparison, but generally can quote little except the "official" indices of product per unit of labour, whose accuracy he has already disproved for the earlier period. He also shows a curious incapacity to bring up to date the statistics of his own country. It is true that a complete Census of Manufactures is not available for the United States since 1947, but an American economist could easily obtain enough information to bring his indices approximately up-to-date. He also confesses his inability to find relevant American prices for certain types of locomotives produced in Russia.

The back-flap of the book, as often happens, quotes a paragraph out of context, which has the effect of slanting it in favour of the now familiar and misleading legend that Russian productivity is steadily increasing at the rate of 6 per cent. per annum.

The figures for the 1930's which he analyses, show, in effect, very erratic rates of growth. The rate is very poor about 1932–3, grows very rapidly in the mid–1930's, then slows down again after 1936. No comment is made on these extraordinary fluctuations. One would have thought that even the most elementary account of Soviet history during that period would have made some reference to the Famine of 1932–3, and the Great Purge which began in 1937. They certainly affected the course of productivity very strongly. Yet there is no mention of them at all.

The Russian product per man in the 1930's, expressed as a percentage of the American (working hours at that time were not very different in the two countries) averages about 40 per cent., the figures ranging from 58 per cent. in tractor manufacture and 54 per cent. in open-cast ore mining to 23 per cent. in cotton weaving and 15 per cent. in the manufacture of heavy construction machinery. The difficulties of comparing products in different countries are discussed, and several references are made to Dr. Snow's eloquent objection to the idea of such comparisons in this *Journal* in 1944.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Professor Galenson shows extremely high rates of increase of productivity in mining and iron and steel, and a high rate (5·1 per cent. per year) in sugar beet. For cotton textiles he shows a figure of 3·9 per cent. per year, for engineering 3·5 per cent. per year and for the shoe industry an actual decline of 1·4 per cent. per year. Many casual readers will no doubt fall into the trap of giving quite undue weight to the experience of a limited number of heavy industries, which have received extremely favourable treatment in the matter of capital investment, and about which (understandably) more abundant supplies of statistical information are forthcoming, and so estimating a high overall average growth rate, which is certainly not applicable to Soviet industry as a whole.

Finally, one very interesting point may be noted. It is a commonplace that, as production becomes more highly mechanized and uses more complex and valuable equipment and components, the ratio of administrative and clerical to direct production workers is bound to rise. This ratio in the United States turns out to be slightly *lower* than it is in Russia.

COLIN CLARK.

14.—*Control Chart Technique when Manufacturing to a Specification.* British Standard 2564 : 1955. By B. P. Dudding and W. J. Jennett. British Standards Institution, 1955. 77 pp. 8½″. 10*s*. 6*d*.

This is, word for word, a re-publication of a booklet with the same title first published by the G.E.C. in 1944. The Units and Standards Committee of the British Standards Institution have now issued it as a British Standard to supplement their other publications on the subject, B.S.600, B.S.600R, B.S.1008 and B.S.1313.

Although B.S.2564 is complete in itself, B.S.600R, with its rather longer introduction to the basic ideas behind the control chart technique, should be read first; especially by those unfamiliar with the subject. As the title indicates, the new Standard deals with the situations which arise when different relations subsist between the control limits of a process and the specification limits; between what the process is doing and what it is required to do. The different cases are identified by a low, medium or high relative precision index which is a comparison of the variability of the process with the variability permitted by specification.

Precise instructions are given for constructing and interpreting charts and a schedule displays the kind of action necessary in each of five cases. The same sub-headings are then repeated in a brief discussion section with examples and reproductions of charts. The whole of this layout is repeated for charts used for counting defectives (gauging instead of measuring). Appendices give definitions of terms used, and tables of the various factors used in calculating limits from average range. Statisticians will wish to note that the factors D, Table 3, are based on an early approximation to the distribution of range and only agree in first decimal with values calculated from more recent tables (*Biometrika Tables*, Vol. 1, Table 23). Appendices C, D and E are a useful collection of miscellaneous notes which might well have formed a main section of the book itself, since they answer questions which are frequently asked about, for example, frequency and method of sampling and sensitivity of the charts.

Although there are now many texts on statistical quality control there are only a few of the manual type known to the reviewer (excluding, of course, those privately issued by firms for their own use). This one is certainly worthy of the wider circulation it now receives; the significance lies in the lapse of eleven years. Are we at last about to see in British Industry a widespread interest in quality and the techniques of controlling it? Is the B.S.I. in front of Industry or behind? Has there been no experience during these eleven years to cause the authors to modify their presentation? There has certainly been some theoretical work, yet, for example, there is no mention of the paper (Keen and Page, *Applied Statistics*, Vol. 2, p. 13) emanating from the same (G.E.C.) laboratory on the use of successive differences instead of range; multiple defects present a problem which is not mentioned at all. Stafford Beer's paper (*Applied Statistics*, Vol. 2, p. 160) on multiple charts has some relevance. Stevens' paper on Control by Gauging (this Journal, Series B, 1948, 10.54) might modify the ideas presented on "10 per cent. inspection gauges". Although the authors quite rightly stress throughout the importance of controlling the process rather than just requiring conformity with specification, their treatment of modified limits which of course forms quite a large section of the book departs from this without specifically saying so. Modified limits like the 10 per cent. inspection gauge are a sop to production and need great care in use.

The B.S.I. itself has through its committees come to realize the value of the statistical quality control technique in establishing the assurance it requires for its Mark scheme, and otherwise ensuring conformity with specification of goods that are stated to conform to a British Standard. There is no mention in this booklet of this aspect of statistical quality control, though the title

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

might lead one to expect it. It might be a good idea for the Institution to assemble and publish its ideas on this and so give a lead for its committees.

E. D. VAN REST.

References

J. KEEN and D. PAGE, *Applied Statistics*—**2** : 13.
STAFFORD BEER, *Applied Statistics*—**2** : 160.
W. L. STEVENS, *J. R. Statist. Soc.*, B, 1948. 10: 54.
B.S.600, *Statistical Methods in Industry*.
B.S. 600R, *Quality Control Charts*.
B.S.1008, *Guide to Quality Control*.
B.S.1313, *Fraction Defective Charts for Quality Control*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL NOTES

(1) British Official Statistics

The interim index of retail prices published by the Ministry of Labour and National Service which was 150 in July, 1955 (prices at June 17th, 1947 = 100) fell to 149 in August and then rose to 150 in September, 152 in October and 154 in November and December. In August there were substantial reductions in the prices of potatoes and cooking apples, partly offset by increases in the prices of eggs, bacon and coal. In September prices rose for beef, pork and bacon, coal and gas. The marked rise in the index in October resulted from higher prices for many food items, including butter, cheese, eggs, and pork and for furniture and some newspapers. There were further increases in November for food items and for coal, electricity, furniture, cigarettes and some articles of clothing. Separate figures on the same basis (i.e. June, 1947 = 100) are not available for the different groups of commodities, but the following are the figures based on January 15th, 1952 = 100.

Date	*Food*	*Rent and Rates*	*Clothing*	*Fuel and Light*	*Household Durable Goods*	*Miscellaneous Goods*	*Services*	*Alcoholic Drinks*	*Tobacco*	*All Items*
Weights:	399	72	98	66	62	44	91	78	90	1,000
July 12th, 1955	125·6	115·0	96·3	113·0	95·8	101·5	115·5	102·8	100·7	113·3
Aug. 16th, ,,	122·1	115·2	96·4	119·2	95·9	101·5	115·7	102·8	100·7	112·5
Sept. 13th, ,,	123·6	115·4	96·4	120·9	96·4	102·3	116·6	103·0	100·7	113·2
Oct. 18th, ,,	125·9	116·4	96·5	121·6	97·6	103·7	116·8	103·1	100·7	114·4
Nov. 15th, ,,	128·0	116·6	97·4	125·7	100·9	106·2	117·1	103·1	102·9	116·2
Dec. 13th, ,,	127·1	116·7	98·3	127·5	101·6	106·5	117·6	103·1	102·9	116·2

The Ministry of Labour index of weekly wage rates, which was 153 in July (June 30th, 1947 = 100) remained at that figure till November when it rose to 154. The principal bodies of workpeople whose rates of wages were raised during the period included iron and steel workers, operatives engaged in wholesale bespoke tailoring and shirt making, textile bleachers and dyers, bus workers and Post Office manipulative staff.

The six-monthly inquiry into average earnings showed that in April, 1955, the average earnings per week in the industries covered by the inquiries were 217*s.*5*d.* for men, 112*s.*5*d.* for women, 90*s.*11*d.* for youths and boys, and 71s.6d. for girls. These figures represent increases over October, 1938, of 215, 246, 249, and 287 per cent. For all workers combined the average was 182*s.* 3*d.*, an increase of 242 per cent. The index for April 1955, compared with 100 at April 1947, was 176. This compares with the weekly wage *rate* index of 152. The average hours worked per week in April, 1955, were 46·9 (48·9 for men only). In manufacturing industries alone the average weekly earnings were 224*s.*7*d.* for men and 112*s.*11*d.* for women. The corresponding hourly earnings were 55·5d. and 32·4d.

The total working population and the numbers in civil employment in the six months ended November, 1955, were as follows:

(*Thousands*)

Date	*Total Working Population*			*Numbers in Civil Employment*		
	Males	*Females*	*Total*	*Males*	*Females*	*Total*
June, 1955 . .	16,064	7,811	23,875	15,162	7,734	22,896
July, ,, . .	16,077	7,839	23,916	15,178	7,760	22,938
Aug., ,, . .	16,128	7,879	24,007	15,230	7,796	23,026
Sept., ,, . .	16,125	7,893	24,018	15,225	7,805	23,030
Oct., ,, . .	16,142	7,929	24,071	15,234	7,835	23,069
Nov., ,, . .	16,112	7,973	24,105	15,225	7,881	23,106

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The numbers of persons on the unemployment registers of Employment Exchanges in Great Britain rose by 41,000 between July and November and fell by 10,000 in the following month. The total of 215,650 in December represented 1.0 per cent. of the estimated total of employees. The percentages in the different Regions ranged from 0.4 in the Midlands to 1.7 in Wales and 2.3 in Scotland.

The following is a sex analysis of the figures:—

Number of Unemployed Persons on the Registers of Employment Exchanges

Date	Men and Boys	Women and Girls	Total
July 11th, 1955 . . .	116,690	68,239	184,929
Aug. 15th, ,, . . .	124,835	74,858	199,693
Sept. 12th, ,, . . .	124,319	77,902	202,221
Oct. 10th, ,, . . .	133,048	82,107	215,155
Nov. 14th, ,, . . .	141,955	83,958	225,913
Dec. 12th, ,, . . .	139,006	76,644	215,650

Of the total of 215,650 in December 60,915 had been unemployed for not more than two weeks, 60,788 from two to eight weeks, and 84,204 for over eight weeks, while 9,743 were temporarily stopped.

The numbers of persons absent from work owing to sickness, including self-employed as well as employed, were 811,400 in July, 798,400 in August, 839,800 in September, 896,500 in October, 910,400 in November and 860,900 in December. The numbers of employed persons absent owing to industrial injuries were 59,100, 53,900, 59,900, 61,100, 60,900, and 55,300.

The aggregate membership of trade unions in the United Kingdom as shown in the *Ministry of Labour Gazette* for October, 1955, was 9,495,000 at the end of 1954. This was 36,000 more than at the end of the previous year. Of the total membership 67 per cent. were in 17 unions with a membership of 100,000 or more, and 11.9 per cent were in unions with a membership of 50,000 to 100,000. General labour organisations accounted for 2,078,000 members, 1,803,000 were in metal trade unions, and 839,000 in coal mining unions.

The Text Volume of the *Registrar General's Statistical Review of England and Wales* for the year 1952 has recently been issued (H.M.S.O. 1955, 8*s.* 0*d.*). This is a commentary on the Tables of medical and vital statistics which have been published earlier; it reviews major statistical trends in fertility, and mortality in England and Wales up to 1952, and makes some reference to other aspects of the work of the General Register Office.

The 344 tables in the latest edition of the *Annual Abstract of Statistics* (No. 92, H.M.S.O., 1955, 21*s.*) mostly cover the post-war years to 1954, with 1938 for a pre-war comparison. They provide comparative statistics of changing social conditions in the United Kingdom and their range is wider than that of the *Monthly Statistical Digest*. Among the new tables are No. 6, "Aliens registered in the U.K."; Nos. 58 to 60, "National Health Service. Summary of main Services Provided"; No. 67, "Comparison of Criminal Charges proved in England and Wales and Scotland" (which is introduced on a recommendation of the Interdepartmental Committee on Social and Economic Research); and No. 207, "Estimated Farming Net Income".

The fourth *Annual Report of the National Food Survey Committee* (H.M.S.O. 1955, 4*s.*) covers conditions in 1953, differentiating for the first time current food consumption and expenditure on a regional basis. Differences observed in 1952 and 1953 between urban and rural households are discussed and a special section describes the diets of the Scottish sample.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(2) Other Statistics

A report has been received from the Indian Statistical Institute, Calcutta, on *The Facilities for Training and Research in Statistics in India*. The report has been compiled by P. B. Patnaik, and is based upon a survey of training and research facilities in universities and other institutions which was carried through in the summer of 1955, at the instance of the Central Statistical Organization of the Government of India and the Indian Statistical Institute. The intention is to launch a further programme of training and research in statistics in connection with the second Five-year Plan, and the first thing was to obtain factual information on the present position. The appendices to the report list: (A) the items of information requested, (B) institutions equipped to provide practical training, (C) main lines of research conducted in universities, (D and following) detailed information as to statistical departments and courses at various institutions. The main body of the report consists of twelve chapters. Undergraduate and post-graduate teaching in statistics is described, also cases where statistics figures as an ancillary subject in other courses. Examinations for diplomas and certificates are detailed, and research facilities in Universities are commented on. Matters of interest in relation to the establishment of departments of statistics are described in relation to universities which have such departments. Chapters follow on technological and research institutions, the Indian Statistical Institute and other institutions.

The main conclusions arising from the survey are as follows. There are both university courses and special institution courses; the former have cultural and scholastic, but only limited professional, objectives, while the latter have definite professional objectives. There are quite a number of universities teaching statistics, but only a small number of special institutions, with limited capacity for training. A certain amount of specialization takes place at some of the universities, but few universities can afford to have real specialists on their staffs. The main professional institutions are the Indian Council of Agricultural Research at New Delhi and the Indian Statistical Institute in Calcutta. Research in universities is mainly theoretical, and funds are needed to extend research on the applied side and to provide for visits of experts from abroad. It is important in connection with the second Five-year Plan to formulate a comprehensive programme of further facilities for teaching and research in statistics, and to attract funds to enable such a programme to be carried out.

The report is critically examined in the *Calcutta Statistical Association Bulletin* for July 1955 (Vol. 6, p. 49).

The new edition of the *Guide to New Zealand Official Statistics* (Christchurch, etc., London, Whitcombe & Tombs, 1955) is more than a list of sources and outline of methods. Where appropriate, chapters are prefaced by an historical and geographical introduction, so that the publication becomes an economic and statistical handbook for the country. The unusual index, by "topics" does not seem to be as helpful as the normal form in leading the enquirer direct to specific information.

CURRENT NOTE

The Fellows named below were included in the New Year's Honours List, 1956:

K.B.E.	Sir Hugh Beaver
Knight Bachelor	Ralph G. Hawtrey, C.B.
C.B.E.	A. H. Marshall
C.B.E.	W. Bertram Nelson

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

OBITUARY

H. W. METHORST—1868–1955

Dr. H. W. Methorst, who died last July at the age of 87, was born at Amersfoot (Holland). He became a student at the University of Utrecht, where he gained his doctorate *en droit* "cum laude" in 1895 in virtue of a thesis entitled "Assistance par le travail aux indigents". At that date he entered the Dutch official statistical service, and rose with its development till in 1927 he was Director-General of the Central Bureau of Statistics, and held that office till his retirement in 1939.

In 1920, recognized as the most prominent statistician in Holland, he was elected an Honorary Fellow of the Royal Statistical Society. Except in connection with the International Institute of Statistics at its Jubilee meeting at the Royal Statistical Society's Centenary in 1934, it does not appear that he had any direct intercourse with the Society.

Outside his own country Methorst was best known as the Secretary General of the International Institute of Statistics. Elected a Fellow in 1907, he was nominated Secretary General in 1913, succeeding his compatriot Verrijn Stuart, who had followed Major Craigie in 1907. Prior to 1913 the Institute had no headquarters. Its business had been transacted at its biennial Sessions in different capital cities. The President and Vice-Presidents were from separate countries, the successive treasurers were English, and the Secretary General was Luigi Bodio from its foundation in 1885 till he became President in 1907. The only publication of the Institute was of the proceedings and papers communicated at its general Sessions.

After deliberations extending over many years, (as described in *50 Années de l'Institut International de Statistique*, 1934, by F. Zahn, President), it was decided in 1913 to establish a Permanent Office of Statistics under the authority of the Institute. Its duties were to aim at improving the comparability of national statistics, publish international data, and prepare for the Institute's regular Sessions reports on the development of scientific statistics.

Owing to the interest and generous support of the Dutch Government and the City of the Hague, due in large measure to Methorst's influence, it was decided that the Permanent Office should be located at the Hague, under direction of the Secretary General of the Institute. Methorst held the offices of Director General of Statistics in the Netherlands, of Secretary General of the International Institute and Director of the Permanent Office, till he resigned in 1947 at the age of nearly 80. Largely owing to his influence and activity the Institute and the Permanent Office obtained considerable subventions and subscriptions from several Governments, especially the Dutch, but not from that of the United Kingdom. It strengthened its position and developed its work during the period 1914–23, during which there were no general Sessions of the Institute owing to the War and the difficulty of re-establishing international meetings. At the Brussels Session (1923) the success and importance of Methorst's work was the subject of very warm appreciation.

During the decade 1923–33 the Institute met at Rome, Cairo, Warsaw, Tokio, Madrid and Mexico. Meanwhile the Permanent Office issued frequent monographs on the organization of statistics in different countries and especially on demographical reports. These were continued, together with original articles by Fellows of the Institute, in the well-known *Revue de l'Institut International de Statistique* initiated in 1933, and now in its twenty-third year.

The second World War, including the invasion of Holland, disorganized the Institute even more than the first. At last, largely under the stimulus of W. F. Willcox of Cornell University, a meeting took place at Washington in 1947 and the Institute has from that date developed with renewed and wider activities. At his advanced age Methorst did not go to Washington and his colleague at the Hague, J. Tinbergen, was elected in his place.

As Treasurer of the Institute, I took part in many meetings with him at the Hague, and travelled with him to Mexico and elsewhere. It was a pleasure to watch the quiet efficiency with which he organized the business and worked with colleagues of different nationalities and with divergent views.

It would be a mistake, however, to think that Methorst was concerned only with organization, whether as a civil servant or the Institute's secretary. As a statistician he was interested especially in the problems of population and health. His last unofficial publication was entitled *Naissances, prosperité et mortalité infantile dans quatres villes et 103 communes rurales* (1935).

A. L. BOWLEY.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Charles Stewart Orwin

Dr. Charles Stewart Orwin, who died on June 30th of last year, was born at Warnham, Sussex, in 1876. He went first to Dulwich College and then entered Wye College in 1895, the second year of its existence as an agricultural institution; the Principal at that time was A. D. Hall, later Sir Daniel Hall, whose work and teaching were destined to influence Orwin's outlook and career to a marked extent. He left in 1897 but after two short periods, first with an estate agent in Tonbridge and then in business in the City, he returned to Wye for three years as lecturer in Estate Management. In 1906, however, he was invited by Mr. Christopher Turnor to manage and develop his Lincolnshire estates and it was his experience in this work during the next seven years which formed the foundation of so much of his later work and which stimulated that essentially practical approach to the agricultural and economic problems to which he turned his attention throughout his life.

Meantime Hall was developing his ideas on agricultural education, one of which was the need for quantitative assessment of farming problems if both teaching and research were to be undertaken effectively. Orwin set to work in Lincolnshire to develop methods of book-keeping and cost-accounting suited to the complexities of British farming. There was nothing pedagogic in the results; he knew from experience that his system must be simple and practical if farmers were to be persuaded that profit and loss should be the principal yardstick of their business. When, in 1913, the Agricultural Economics Research Institute was set up in the University of Oxford by the Development Commission, Orwin was appointed its first Director, the position which he held until his retirement to Blewbury in 1946.

At first he combined his work at Oxford with help and advice in spheres of Government. His work in farm costings was of direct importance in the First World War and immediately afterwards when attempts were made to control agricultural prices on the basis of cost of production. In 1917 he was made a member of the first Agricultural Wages Board and in 1922 he was appointed Agricultural Assessor to the Tribunal of Investigation which was set up during the severe depression which accompanied the years of deflation after the War; the Tribunal's report, in the preparation of which Orwin played a very considerable part, was a remarkable survey which was written during those most difficult years and which has subsequently stood the test of time. As time went on, however, his work became more centred in the Institute and the University. Governmental work, which involved visits to London, and the shallowness and empiricism of agricultural politics in the inter-War years did not appeal to him.

Soon after his arrival in Oxford, Orwin was made a member of Balliol College Senior Common Room, an association which was a source of great happiness in his life. He became a Fellow in 1922 and subsequently Estates Bursar and finally, in 1946, an Honorary Fellow. Though the College undoubtedly benefited from his devoted service, he himself would have asserted that the advantages were entirely his; through the freemasonry of Common Room life he learned quickly to understand the University and the University gradually learned of his Institute and its work; the management of the College's estates continued his earlier contact with the practical problems of English farming and helped to satisfy that almost passionate love for the English countryside from which sprang his devotion to its problems; finally, Balliol gave him those contacts with other scholars which the catholicity of his interests demanded.

But, apart from his family, his Institute was the centre of his Oxford life. His early years in Oxford were years of struggle—to establish a new type of research institution in a new subject in a University where academic innovation as such had little instantaneous appeal. At the end of his thirty-three years as Director, however, the Institute was a well-established part of the University, its staff were members of the University and it was the recognized national centre for work in agricultural economics. This was not a position which developed automatically; it was the result of Orwin's own personality, patience and persistence.

His appointment as the Institute's first Director was a singularly happy one; he had practically all the qualifications for which this new work called. He was always ready to listen to new ideas and the possibilities of new lines of work; he gathered round him young people and gave them every opportunity to prove themselves and their work. There developed in the Institute a strong department in farm management based at first on Orwin's system of cost-accounting but later on newer techniques such as enterprise studies and the survey method. Work on rural industries and agricultural marketing was commenced early in the 'twenties, followed later by studies on land tenure and agricultural credit. In the 'thirties the Institute's scope was extended to cover work in supplies and prices. A list of the Institute's publications during his years as Director reflects the width of his interests and testifies to the freedom and scope which he allowed to his

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

OBITUARY

H. W. METHORST—1868–1955

Dr. H. W. Methorst, who died last July at the age of 87, was born at Amersfoot (Holland). He became a student at the University of Utrecht, where he gained his doctorate *en droit* "cum laude" in 1895 in virtue of a thesis entitled "Assistance par le travail aux indigents". At that date he entered the Dutch official statistical service, and rose with its development till in 1927 he was Director-General of the Central Bureau of Statistics, and held that office till his retirement in 1939.

In 1920, recognized as the most prominent statistician in Holland, he was elected an Honorary Fellow of the Royal Statistical Society. Except in connection with the International Institute of Statistics at its Jubilee meeting at the Royal Statistical Society's Centenary in 1934, it does not appear that he had any direct intercourse with the Society.

Outside his own country Methorst was best known as the Secretary General of the International Institute of Statistics. Elected a Fellow in 1907, he was nominated Secretary General in 1913, succeeding his compatriot Verrijn Stuart, who had followed Major Craigie in 1907. Prior to 1913 the Institute had no headquarters. Its business had been transacted at its biennial Sessions in different capital cities. The President and Vice-Presidents were from separate countries, the successive treasurers were English, and the Secretary General was Luigi Bodio from its foundation in 1885 till he became President in 1907. The only publication of the Institute was of the proceedings and papers communicated at its general Sessions.

After deliberations extending over many years, (as described in *50 Années de l'Institut International de Statistique*, 1934, by F. Zahn, President), it was decided in 1913 to establish a Permanent Office of Statistics under the authority of the Institute. Its duties were to aim at improving the comparability of national statistics, publish international data, and prepare for the Institute's regular Sessions reports on the development of scientific statistics.

Owing to the interest and generous support of the Dutch Government and the City of the Hague, due in large measure to Methorst's influence, it was decided that the Permanent Office should be located at the Hague, under direction of the Secretary General of the Institute. Methorst held the offices of Director General of Statistics in the Netherlands, of Secretary General of the International Institute and Director of the Permanent Office, till he resigned in 1947 at the age of nearly 80. Largely owing to his influence and activity the Institute and the Permanent Office obtained considerable subventions and subscriptions from several Governments, especially the Dutch, but not from that of the United Kingdom. It strengthened its position and developed its work during the period 1914–23, during which there were no general Sessions of the Institute owing to the War and the difficulty of re-establishing international meetings. At the Brussels Session (1923) the success and importance of Methorst's work was the subject of very warm appreciation.

During the decade 1923–33 the Institute met at Rome, Cairo, Warsaw, Tokio, Madrid and Mexico. Meanwhile the Permanent Office issued frequent monographs on the organization of statistics in different countries and especially on demographical reports. These were continued, together with original articles by Fellows of the Institute, in the well-known *Revue de l'Institut International de Statistique* initiated in 1933, and now in its twenty-third year.

The second World War, including the invasion of Holland, disorganized the Institute even more than the first. At last, largely under the stimulus of W. F. Willcox of Cornell University, a meeting took place at Washington in 1947 and the Institute has from that date developed with renewed and wider activities. At his advanced age Methorst did not go to Washington and his colleague at the Hague, J. Tinbergen, was elected in his place.

As Treasurer of the Institute, I took part in many meetings with him at the Hague, and travelled with him to Mexico and elsewhere. It was a pleasure to watch the quiet efficiency with which he organized the business and worked with colleagues of different nationalities and with divergent views.

It would be a mistake, however, to think that Methorst was concerned only with organization, whether as a civil servant or the Institute's secretary. As a statistician he was interested especially in the problems of population and health. His last unofficial publication was entitled *Naissances, prosperité et mortalité infantile dans quatres villes et 103 communes rurales* (1935).

A. L. BOWLEY.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Charles Stewart Orwin

Dr. Charles Stewart Orwin, who died on June 30th of last year, was born at Warnham, Sussex, in 1876. He went first to Dulwich College and then entered Wye College in 1895, the second year of its existence as an agricultural institution; the Principal at that time was A. D. Hall, later Sir Daniel Hall, whose work and teaching were destined to influence Orwin's outlook and career to a marked extent. He left in 1897 but after two short periods, first with an estate agent in Tonbridge and then in business in the City, he returned to Wye for three years as lecturer in Estate Management. In 1906, however, he was invited by Mr. Christopher Turnor to manage and develop his Lincolnshire estates and it was his experience in this work during the next seven years which formed the foundation of so much of his later work and which stimulated that essentially practical approach to the agricultural and economic problems to which he turned his attention throughout his life.

Meantime Hall was developing his ideas on agricultural education, one of which was the need for quantitative assessment of farming problems if both teaching and research were to be undertaken effectively. Orwin set to work in Lincolnshire to develop methods of book-keeping and cost-accounting suited to the complexities of British farming. There was nothing pedagogic in the results; he knew from experience that his system must be simple and practical if farmers were to be persuaded that profit and loss should be the principal yardstick of their business. When, in 1913, the Agricultural Economics Research Institute was set up in the University of Oxford by the Development Commission, Orwin was appointed its first Director, the position which he held until his retirement to Blewbury in 1946.

At first he combined his work at Oxford with help and advice in spheres of Government. His work in farm costings was of direct importance in the First World War and immediately afterwards when attempts were made to control agricultural prices on the basis of cost of production. In 1917 he was made a member of the first Agricultural Wages Board and in 1922 he was appointed Agricultural Assessor to the Tribunal of Investigation which was set up during the severe depression which accompanied the years of deflation after the War; the Tribunal's report, in the preparation of which Orwin played a very considerable part, was a remarkable survey which was written during those most difficult years and which has subsequently stood the test of time. As time went on, however, his work became more centred in the Institute and the University. Governmental work, which involved visits to London, and the shallowness and empiricism of agricultural politics in the inter-War years did not appeal to him.

Soon after his arrival in Oxford, Orwin was made a member of Balliol College Senior Common Room, an association which was a source of great happiness in his life. He became a Fellow in 1922 and subsequently Estates Bursar and finally, in 1946, an Honorary Fellow. Though the College undoubtedly benefited from his devoted service, he himself would have asserted that the advantages were entirely his; through the freemasonry of Common Room life he learned quickly to understand the University and the University gradually learned of his Institute and its work; the management of the College's estates continued his earlier contact with the practical problems of English farming and helped to satisfy that almost passionate love for the English countryside from which sprang his devotion to its problems; finally, Balliol gave him those contacts with other scholars which the catholicity of his interests demanded.

But, apart from his family, his Institute was the centre of his Oxford life. His early years in Oxford were years of struggle—to establish a new type of research institution in a new subject in a University where academic innovation as such had little instantaneous appeal. At the end of his thirty-three years as Director, however, the Institute was a well-established part of the University, its staff were members of the University and it was the recognized national centre for work in agricultural economics. This was not a position which developed automatically; it was the result of Orwin's own personality, patience and persistence.

His appointment as the Institute's first Director was a singularly happy one; he had practically all the qualifications for which this new work called. He was always ready to listen to new ideas and the possibilities of new lines of work; he gathered round him young people and gave them every opportunity to prove themselves and their work. There developed in the Institute a strong department in farm management based at first on Orwin's system of cost-accounting but later on newer techniques such as enterprise studies and the survey method. Work on rural industries and agricultural marketing was commenced early in the 'twenties, followed later by studies on land tenure and agricultural credit. In the 'thirties the Institute's scope was extended to cover work in supplies and prices. A list of the Institute's publications during his years as Director reflects the width of his interests and testifies to the freedom and scope which he allowed to his

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

colleagues. He was in no sense a "director" of research; a staunch upholder of academic freedom, he preferred to remain in the background, to stimulate, to encourage and to help. His own standard of scholarship was high and his aim was to assist others to attain it. Lucid and succinct in his own style of writing he sought to promote this in others. Above all, he urged writers to remember for whom they were writing; his articles written for practical farmers in the *Farm Economist* were as much models of style and clarity as his works of pure scholarship.

His own work developed in three main stages. There was first his work on cost-accounting which has already been mentioned. This was followed by his studies on land tenure, including his treatise on land nationalization; his contributions here are best illustrated in *The Tenure of Agricultural Land* (with W. R. Peel, 1925), *The Future of Farming* (1930), *Back to the Land* (with W. F. Darke, 1935), and *Speed the Plough* (1942). Finally, there were his studies in agricultural history. In 1929 he published his *The Reclamation of Exmoor Forest*, a beautifully written account of the history of the countryside around Minehead which he loved so well. His contribution to *Johnson's England* (1933) shows his ability to fit his particular work into a wider context. Then, in 1939, there appeared that very important contribution to British history, *The Open Fields*, in which he shared authorship with his second wife. This detailed and penetrating study of Laxton Manor is now the accepted source of factual knowledge on manorial farming and the subsequent enclosures. Covering a wider range is his very readable *History of English Farming* (1949) and an unfinished history covering the period 1840 to 1940, which, it is to be hoped, will be completed and published.

Throughout all these years of academic work he never lost touch with the science of agriculture or with its practice. His editorship of the *Journal of the Royal Agricultural Society* for twenty-five years from 1912 to 1927 was proof of the former; his experience as a land agent and bursar, combined with constant contact with farmers including such pioneers in farming progress as the Baylis brothers, Hosier, Bamford, Higgs and Abbott, assured that he knew the current trends in the industry for which he was working.

While his own direct contributions to knowledge were substantial, his indirect influence has been possibly even greater. Hundreds of post-graduate students, who are now spread throughout the world, were trained in their work at the Institute; practically all the agricultural economists at work in this country passed at some time of their career under his influence. But those who owe him the greatest debt are those who were his colleagues. His independence of thought, his self-discipline and his high standards of scholarhip, his trust in his staff and the licence which he granted to them, his eagerness to assign credit to where it belonged—especially to younger workers—his meticulous care in the presentation of his studies, his courtesy, his gentleness, his consideration for other people's feelings and his charity were a few of his qualities which they will never forget. He inspired enthusiasm, respect and, above all, affection, even among those whose views differed from his. His innate shyness and sensitivity may have masked his real self for those who knew him only slightly; those who knew him well can only count themselves thrice blessed for his friendship, his inspiration and his unfailing help.

KEITH A. H. MURRAY.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL AND ECONOMIC ARTICLES IN RECENT PERIODICALS

UNITED KINGDOM—

Accounting Research—

October 1955—Social accounts and input-output tables: *E. Grasberg*. The measurement of profits—a study in methods. II. The sample of British companies: *R. K. Nigam*. The Royal Commission and depreciation allowances: *D. Walker*.

Annals of Human Genetics—

October 1955—The estimation of gene frequencies in a random-mating population: *R. Ceppellini, M. Siniscalco* and *C. A. B. Smith*. Non-randomness in consanguineous marriage: *N. E. Morton*. The inheritance of human birthweight: *N. E. Morton*. Blood groups of the Northern Nilotes: *D. F. Roberts, E. W. Ikin* and *A. E. Mourant*. Genetics of dermal ridges: frequency distributions of total finger ridge-count: *S. B. Holt*.

Biometrika—

December 1955—Population estimation based on change of composition caused by a selective removal: *D. G. Chapman*. An age-dependent birth and death process: *W. A. O'N. Waugh*. A large-sample bioassay design with random doses and uncertain concentration: *F. C. Andrews* and *H. Chernoff*. An exact test for correlation between time series: *E. J. Hannan*. Serial correlation in regression analysis: *G. S. Watson*. Some theorems and sufficiency conditions for the maximum-likelihood estimator of an unknown parameter in a simple Markov chain: *J. Gani*. Significance tests for discriminant functions and linear functional relationships: *E. J. Williams*. The use of transformations and maximum likelihood in the analysis of quantal experiments involving two treatments: *F. Yates*. A note on the application of the combination of probabilities test to a set of 2×2 tables: *F. Yates*. A test for homogeneity of the marginal distribution in a two-way classification: *A. Stuart*. Distribution of Kendall's tau based on partially ordered systems: *S. Haberman*. On a class of skew distribution functions: *H. A. Simon*. Simultaneous tests of linear hypothesis: *M. N. Ghosh*. Rank analysis of incomplete block designs. III. Some large-sample results on estimation and power for a method of paired comparisons: *R. A. Bradley*. A method of assigning confidence limits to linear combinations of variances: *A. Huitson*. Interpolations and approximations related to the normal range: *J. W. Tukey*. The gambler's ruin problem with correlation: *C. Mohan*. Tables for significance tests of 2×2 contingency tables: *P. Armsen*. A note on moving ranges: *H. A. David* Censored samples from truncated normal distributions: *A. C. Cohen, Jr.* The rapid calculation of χ^2 as a test of homogeneity from a $2 \times n$ table: *J. B. S. Haldane*. The "inefficiency" of the sample median for many familiar symmetric distributions: *J. T. Chu*. A simple method of calculating the exact probability in 2×2 contingency tables with small marginal totals: *P. H. Leslie*. A test for a change in a parameter occurring at an unknown point: *E. S. Page*. A paradox in statistical estimation: *A. Stuart*. Cumulants of a transformed variate: *G. S. James*. The likelihood ratio test for Markoff chains: *I. J. Good*. Exact forms of some invariants for distributions admitting sufficient statistics: *V. S. Huzurbazar*.

British Journal of Preventive & Social Medicine—

July 1955—Factor analysis of the disabilities of the elderly with the aid of an electronic calculator: *J. Maddison, R. C. F. Smith* and *O. Olbrich*. Vital statistics of the county of London in the years 1901 to 1951: *W. J. Martin*. Statistical experimentation therapeutics: *R. Wrighton*. Changing patterns of mortality in England and Wales. III. Mortality in old age: *W. Taylor* and *M. Dauncey*.

October 1955—Oxford Child Health Survey: effect of childish ailments on skeletal development: *D. Hewitt, C. K. Westropp* and *R. M. Acheson*. Changing population of mental hospitals: *G. M. Carstairs, W. L. Tonge, N. O'Connor* and *L. E. D. Barber*. Mental illness and social class in Bristol: *E. H. Hare*. Records of deaf mutism in Northern Ireland: *E. A. Cheeseman* and *A. C. Stevenson*. Measurement and interpretation of subcutaneous fat, with norms for children and young adult males: *W. H. Hammond*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

British Journal of Sociology—

December 1955—Some developments in prediction methodology in applied social research: *L. T. Wilkins.*

British Journal of Statistical Psychology—

November 1955—The determinacy of factor score matrices with implications for five other basic problems of common-factor theory: *L. Guttman.* Factor rotation for proportional profiles: analytical solution and an example: *R. B.* and *A. K. S. Cattell.* Mechanism involved in group pressures on deviate-members: *H. A. Simon* and *H. Guetzkow.* Test reliability estimated by analysis of variance: *C. Burt.* Test reliability in terms of factor theory: *A. F. Mahmoud.*

British Transport Review—

August 1955—"Electronic brains": *R. Glendinning.*

Cambridge Philosophical Society, Proceedings—

October 1955—The use of economized polynomials in mathematical tables: *C. W. Clenshaw* and *F. W. J. Olver.* Extensions of a renewal theorem: *W. L. Smith.* Correlated random walk: *J. Gillis.* Sampling from an isotropic Gaussian process: *J. M. Hammersley* and *J. A. Nelder.* Stochastic differential equations: *D. A. Edwards* and *J. E. Moyal.*

Economica—

November 1955—Competition in the retail grocery trade, 1850–1939: *F. G. Pennance* and *B. S. Yamey.* The purchasing power of British exports further considered: *D. J. Morgan* and *F. W. Paish.* On the calculation of tax burdens: *A. H. Conrad.*

Economic Journal—

September 1955—Convertibility and triangular trade as safeguards against economic depression: *C. M. Wright.* Trade imbalance, gains from trade and national income change: *V. K. Ramaswani.* Industrial production and profits in the United Kingdom and the United States: *E. H. Stern.*

Institute of Actuaries, Journal—

Vol. 81, Part II, No. 358—The components of mortality: *H. A. R. Barnett.* The World Population Conference, Rome, 1954: *B. Benjamin, P. R. Cox* and *F. A. A. Menzler.*

Manchester Schoool of Economic and Social Studies—

September 1955—Econometric models and the average duration of business cycles: *J. Johnston.* National insurance contributions, 1946–1955: *D. Walker.* Owner-farming in England and Wales 1900 to 1950; *S. G. Sturmey.* Some aspects of urban developments by colliery companies, 1919–1939: *P. H. White.*

Oxford Economic Papers (New Series)—

October 1955—Import prices and inflation: the experience of the United Kingdom 1950–2: *G. W. Maynard.* Investment and location in the steel industry: *G. R. Denton.* Gross margins and efficiency measurement in retail trade: *M. Hall* and *J. Knapp.*

Oxford University Institute of Statistics Bulletin—

November 1955—Some problems raised by the Report of the Royal Commission on the Taxation of Profits and Income: *P. P. Streeten.* Output, employment, and productivity in British manufacturing in the last half century: *A. Maddison.* Major consumer expenditure and ownership of durable goods: *L. R. Klein.*

Population Studies—

November 1955—Medical evidence related to English population changes in the eighteenth century: *T. McKeown* and *R. G. Brown.* The supply of women for gainful work in Britain: *C. E. V. Leser.* Fertility differentials in Detroit, 1920–1950: *A. Mayer* and *C. Klapprodt.* Age at marriage and educational attainment in the United States: *C. Tietze* and *P. Lauriat.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Review of Economic Studies—

Vol. XXIII (1) No. 60—Safety first and hedging: *L. G. Telser.* The use of index numbers in demand analysis: *R. Bergstrom.* The Pareto distribution and the Cobb/Douglas production function in demand analysis: *H. S. Houthakker.* A note on block tariffs: *A. Gabor.* Economic planning and the long term programme: *J. Mitchell.* Specialization in production: *L. McKenzie.* Uncertainty and the cobweb: *F. H. Hahn.* Components of a change in the balance of trade: *G. D. N. Worswick.*

AUSTRALIA—

Statistical Society of New South Wales, Bulletin—

June 1955—Observation and experiment in medicine: *H. O. Lancaster.*

INDIA—

Sankhyā—

July 1955—Completeness, similar regions, and unbiased estimation—Part II: *E. L. Lehmann* and *H. Scheffé.* Certain modified forms of binomial and Poisson distributions: *V. M. Dandekar.* A note on the structure of a stochastic model considered by V. M. Dandekar: *D. Basu.* Analysis of dispersion for multiple classified data with unequal numbers in cells: *C. R. Rao.* Approximate probability values for observed number of "successes" from statistically independent binomial events with unequal probabilities: *J. E. Walsh.* Tables of two-sided 5% and 1% control limits for individual observations of the *r-th* order: *M. Masuyama.* Modified mean square successive difference with an exact distribution: *A. R. Kamat.* Estimation and tests of significance of the components of a time-series: *O. Suryanarayana.* A note on a form of Tchebycheff's inequality for two or more variables: *D. N. Lal.* Unbiased test for a specified value of the parameter in the non-central *F*-distribution: *N. Marakathavalli.*

Calcutta Statistical Association Bulletin—

July 1955—Joint test for the mean and variance of a normal population: *P. K. Bhattacharya.* A note on moments of order statistics from symmetrical population: *K. C. Seal.*

UNITED STATES—

American Economic Review—

December 1955—Interregional input-output analysis: *L. N. Moses.* The theory of occupational wage differentials: *M. W. Reder.*

American Statistical Association, Journal—

December 1955—Statistics of the 1954 polio vaccine trials: *K. A. Brownlee.* A study of industrial use of probability statistics in the physical sciences: *J. H. Toulouse.* Time series problems in aeronautics: *H. Press.* Use of the range instead of the standard deviation: *G. E. Noether.* A diagrammatic representation of the sum of squares and products: *C. C. Li.* The analysis of variance: a graphical representation of a statistical concept: *B. A. Barnes, E. Pearson* and *S. Reiss.* Nomograph for computing multiple correlation coefficients: *F. M. Lord.* Applications of correlation models for biserial data: *R. F. Tate.* A multiple comparison procedure for comparing several treatments with a control: *C. W. Dunnett.* The distribution of the quotient of maximum values in samples from a rectangular distribution: *V. N. Murty.* Maximum likelihood estimation of the dispersion parameter of a chi-distributed radial error from truncated and censored samples with applications to target analysis: *A. C. Cohen, Jr.* The distribution of the product of maximum values in samples from a rectangular distribution: *P. R. Rider.* Fixed, mixed, and random models: *M. B. Wilk* and *O. Kempthorne.* Components of a difference between two rates: *E. M. Kitagawa.* Recent developments in the sampling of human populations in Great Britain: *C. A. Moser.* Uses of crude vital rates in the analysis of reproductivity: *G. J. Stolnitz.* Differential mortality, general and cause-specific in Buffalo, 1939–41: *C. Yeracaris.* Interest as a source of personal income and tax revenue: *L. H. Selzer.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

American Statistician—

October 1955—Federal statistical activities. Diffusion indexes: a comment: *G. Moore.*

Annals of Mathematical Statistics—

September 1955—Meyer Abraham Girschick, 1908–1955: *D. Blackwell* and *A. H. Bowker.* The treatment of ties in some non-parametric tests: *J. Potter.* On a class of decision procedures for ranking means of normal populations: *K. C. Seal.* Ordered families of distributions: *E. L. Lehmann.* On an application of Kronecker product of matrices to statistical designs: *M. N. Vartak.* Bounds for the distribution function of a sum of independent, identically distributed random variables: *W. Hoeffding* and *S. S. Shrikhande.* Some minimax invariant procedures for estimating a cumulative distribution function: *Om P. Aggarwal.* Distribution of quadratic forms and some applications: *A. Grad* and *H. Solomon.* Urn models of correlation: *J. A. McFadden.* Statistics and subfields: *R. R. Bahadur.* Estimation of parameters of truncated or censored exponential distributions: *W. L. Deemer, Jr.* and *D. F. Votaw, Jr.* Estimation of the mean and standard deviation by order statistics. Part II: *A. E. Sarhan.* Probability of indecomposability of a random mapping function: *L. Katz.* A necessary and sufficient condition for admissibility: *C. Stein.* A bivariate sign test: *J. L. Hodges, Jr.* On the convergence of empiric distribution functions: *J. R. Blum.* Calculation of exact sampling distribution of ranges from a discrete population: *I. W. Burr.* The stochastic convergence of a function of sample successive differences: *L. Weiss.*

Biometrics—

September 1955—Experimental design in industry: *H. C. Hamaker.* The exploration and exploitation of response surfaces: an example of the link between the fitted surface and the basic mechanism of the system: *G. E. P. Box* and *P. V. Youle.* Design and analysis of two phase experiments: *G. A. McIntyre.* The Bradley-Terry probability model and preference tasting: *N. T. Gridgeman.* Sur la determination de l'axe d'un nuage rectiligne de points: *G. Teissier.* The variance of the genetic correlation coefficient: *E. C. R. Reeve.* Tests for linear trends in proportions and frequencies: *P. Armitage.* An example of the truncated Poisson distribution: *D. J. Finney* and *G. C. Varley.*

Econometrica—

October 1955—Survey of economic forecasting techniques (a Survey article): *C. F. Roos.* Correlation and regression estimates when the data are ratios: *E. Kuh* and *J. R. Meyer.* The variation of technical coefficients in the antifriction bearing industry: *A. Phillips.* Production functions and British coal mining: *C. E. V. Leser.* Note on an inventory problem discussed by Modigliani and Hohn: *F. Morin.*

Estadística—

Next tasks in the measurement of production and productivity: *I. H. Seigel.* Estadísticas migratorias internacionales: situación actual, aplicación y futuro desarrollo de las recomendaciones internacionales: *Oficina de Estadística de las Naciones Unidas.* Procedimiento de Interpolacion para la construccion de una tabla completa de mortalidad: *R. Garica-Frias.* Procedimientos de Recopilación y clases de datos de las estadísticas industriales continuas en Argentina: *Direccion Nacional del Servicio Estadistico.* Subsídios para o Estudo dos Fatores Determinantes e Consequências das Variações Demográficas no Brasil: *G. Jardim.* Vital statistics of Panama: *C. A. Miró.*

Harvard Business Review—

September–October 1955—Decision making in the age of automation: *M. L. Hurni.* "Clocks" for management control: *A. W. Rucker.* How to figure equipment replacement: *P. A. Scheuble, Jr.* The employee performance record: a vew appraisal and development tool: *J. C. Flanagan* and *R. K. Burns.*

Industrial Quality Control—

October 1955—Industrial applications of operations research: *J. B. Lathrop.* OR concepts useful in QC: *H. O. Davidson.* QC concepts useful in OR: *P. S. Olmstead.* The master control system in General Electric: *N. P. Demos.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

November 1955—Statistical technique for forecasting sales: *G. G. Parkin.* Quality control in paper finishing: *E. R. Hoffman.* Statistical design in electronics production-line experimentation: *F. Caplan, Jr.* Statistical design of experiments in metallurgical research: *S. Gilbert.* Statistical principles of testing: *J. Mandel.*

December 1955—Continuous process control in a petroleum refinery: *J. T. Walter.* The application of quality control techniques in determining work assignments and standards: *S. Barkin.* An engineer evaluates statistical methods: *N. Coda.* Textile quality control: *C. C. Wilson.*

Journal of Political Economy—

December 1955—The London *Economist* and the high tide of laissez faire: *S. Gorson.*

Mathematical Tables and Other Aids to Computation—

July 1955—Numerical calculation of certain definite integrals by Poisson's summation formula: *H. E. Fettis.* Trapezoidal methods of approximating solutions of differential equations: *P. C. Hammer* and *J. W. Hollingsworth.* Solving systems of linear equations with a positive, definite, symmetric, but possibly ill-conditioned matrix: *J. D. Riley.*

Milbank Memorial Fund Quarterly—

October 1955—Reading skills and school entrance age: *E. M. Gruenberg* and *H. Birch.* Maternal and newborn nutrition studies at Philadelphia Lying-In Hospital. Newborn Studies. III. Size and growth of babies during the first three months of life: *R. V. Kasius, A. Randall, W. T. Tompkins* and *D. G. Wiehl.* Changes in fertility by socio-economic status during 1940–1950: *C. V. Kiser.*

Operations Research Society of America, Journal—

August 1955—From Plato to linear program: *Earl of Halsbury.* Stochastic properties of waiting lines: *P. M. Morse.* An analysis of a railroad classification yard: *R. R. Crane, F. B. Brown* and *R. O. Blanchard.* Application of discrete variable servo theory to inventory control: *H. J. Vassian.* Road safety and traffic research in Great Britain: *W. H. Glanville.* Production scheduling for an arbitrary number of periods given the sales forecast in the form of a probability distribution: *G. Dannerstedt.* Production and inventory control in a chemical process: *R. L. Ackoff.*

November 1956—A challenge to operations research: *G. L. Parkhurst.* An analog solution of programming problems: *C. M. Arlow* and *G. Brigham.* Parametric objective function (part 2)—generalization: *S. I. Gass* and *T. L. Saaty.* Applications of game theory in fighter versus bomber combat: *T. E. Caywood* and *C. J. Thomas.* On a congestion problem in an aircraft factory: *G. Brigham.* A non-numerical approach to production scheduling problems: *S. B. Akers* and *J. Friedman.* Determination of the maximal steady state flow of traffic through a railroad network: *A. W. Boldyreff.* An appreciation of systems analysis: *C. Hitch.* Cost-effectiveness analysis for strategic decisions: *R. B. Foster* and *F. P. Hoeber.* Symposium on applications of operations research to urban services. 1. A determination of the optimum location of fire-fighting units in New York City: *D. Valinsky*; 2. Expectancy of multiple vehicular breakdowns in a tunnel: *L. C. Edie*; 3. Evaluating the adequacy of airport parking lots: *F. V. Hurst, Jr.*; 4. Engineering studies of urban traffic flow: *H. A. Barnes.*

Psychometrica—

September 1955—A generalized simplex for factor analysis: *L. Guttman.* Equating test scores—a maximum likelihood solution: *F. M. Lord.* Axioms of a theory of discriminating learning: *F. Restle.* The objective definition of simple structure in linear factor analysis: *L. R. Tucker.* *F*-test bias for experimental designs in educational research: *N. Gourlay.* Least squares estimates and optimal classification: *H. E. Brogden.* An improved method for tetrachoric *r*: *W. L. Jenkins.*

Quarterly Journal of Economics—

August 1955—Investment criteria, productivity, and economic development: *W. Galenson* and *H. Leibenstein.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Review of Economics and Statistics—

August 1955—Acceleration and related theories of investment: an empirical inquiry: *J. Meyer* and *E. Kuh*. Direct effects of a United States recession on imports: expectations and events: *H. K. Zassenhaus*. On the stability of consumer expectations: *R. Ferber*. Index of seasonal variation in unemployment: *W. L. Hansen*.

November 1955—National income, 1954 edition: *T. C. Schelling*. Some notes on the Sterling Area's postwar balance of payments: *R. W. Stevens*. Regional analysis: an inter-industry model of Utah: *F. T. Moore* and *J. W. Petersen*. An inter-city comparison of differentials in earnings and the city workers' cost of living: *J. P. Henderson*. Japan's trade position in a changing world market: *K. K. Kurihara*. Community size, region, labor force, and income, 1950: *E. Mansfield*.

AUSTRIA—

Statistische Vierteljahreschrift—

Vol. VIII, Part 2—Zur Bewertung von Mengenbeständen und Mengenbewegungen: *A. Adam*. Planungsforschung (operational research): *W. Eberl*. Eine Stichprobenerhebung in Wien im Jahre 1859: *W. Horak*.

BRAZIL—

Revista Brasileira de Estatistica—

April–June 1955—Estimativas de máxima verossimilhança com determinação não independentes e acumuladas: *E. Amaral*.

DENMARK—

Nationaløkonomisk Tidsskrift—

Vol. 93, Parts 3–4—Linear programming i produktionsteorien. I: *S. Danø*.

FRANCE—

Revue d'Économie Politique—

July–October 1955—La France économique en 1954 (whole number).

Journal de la Société Statistique de Paris—

April–June 1955—Le statisticien, cet artiste, ou de quelques définitions du mot "Statistique": *M. Dumas*. Méthode d'analyse de la pyramide des ressources et de la structure du pouvoir d'achat: *P. Vinot*.

GERMANY—

Mitteilungsblatt für Mathematische Statistik—

Vol. 7, Part 2—Eine "nicht-parametrische" Ableitung der Streuung: *O. Anderson*. Zur Ermittlung der Abgangslinie für wirtschaftliche und technische Gesamtheiten: *K. Stange*. Einige Bemerkungen zur logarithmischen Normalverteilung: *R. Wartmann*.

Vol. 7, Part 3—Statistische Modelle als Hilfsmittel der Naturbeschreibung: *H. v. Schelling*. Die ausgelesene Normalverteilung im Wahrscheinlichkeitsnetz: *K. Stange*. Über die innere Oberfläche eines Haufwerks von Kugeln: *H. L. de Vries*.

Weltwirtschaftliches Archiv—

Vol. 75, Part 1—Die generative Struktur von Bevölkerungen und Sozialschichten: *G. Mackenroth*. Grundbegriffe der Produktionstheorie vom Standpunkt der Aktivitätsanalyse: *M. J. Beckmann*.

Vol. 75, Part 2—Eine analyse des Vollkostenprinzips: *K. C. Kuhlo*. Long term perspectives of the dollar problem: *T. Surányi-Unger*. The structure of competition in the fertilizer industries in North America: *M. Lamer*.

NETHERLANDS—

Statistica—

Vol. 9, No. 3—Het voorspellen van levertijden uit gegevens over de voortgang van een project: *H. L. Muller*. De intelligentie te Amsterdam en verband met demografische en socio-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

logische kenmerken van de bevolking: *P. de Woolf* and *J. Meerdink*. Confidence intervals associated with a straight line fitted by least squares: *J. E. Kerrich*. Een toets tegen verloop voor een aantal kansen: *C. van Eeden*. Statistische aspecten der muziek: *L. Yntema*.

ITALY—

Statistica—

July–September 1955—Ricerche sulla distribuzione dei redditi familiari nelle categorie economico-sociali del Comune di Ferrara: *A. Bellettini*. Intorno al concetto statistico di leva del lavoro: *R. Lenzi*. Table of inverted matrices for the solution of quadratic regression coefficients: *J. Quartey*. Nota alla tabella per il calcolo dei coefficienti di regressione quadratica: *D. Fürst*. Ancora sul problema del rendimento dei film in Italia: *L. Bergonzini*.

October–December 1955—Fattori sociali ed omogeneizzazione nel comportamento demografico: *N. Federici*. Rilievi statistici sulla frequenza della attività rapida dell'elettroencefalogramma umano normale: *G. B-M. Nistri*. Semplici procedimenti di interpolazione di una funzione esponenziale doppia (catenaria): *W. Capecchi*. L'influenza di alcuni fattori tecnici sul rendimento dei film: *A. Bellettini*.

SPAIN—

Trabajos de Estadistica—

Vol. VI, Part 1—Problemas de máximos y minimos, relacionados con la inferencia en poblaciones finitas: *S. Rios*. The cumulants of a sample mean from a finite population of first N. integers: *B. M. Bennett*. La Estadística: *R. A. Fisher*. Sobrevisión por muestro en las Universidades españolas: *F. Azorin*. Notas para la enseñanza de algunos conceptos elementales de Estadística en el Bachillerato: *E. G. Espana*.

SWEDEN—

Ekonomisk Tidskrift—

December 1955—Expectation, income and profit: *G. L. S. Shackle*.

Skandinavisk Aktuarietidskrift—

Parts 1–2 1955—A form of Neyman's $\Psi^2\kappa$ test of goodness of fit applicable to grouped and discrete data: *D. E. Barton*. On the cumulants of the logarithmic generalized variance and variance ratio: *B. M. Bennett*. When does ruin occur in the collective theory of risk?: *C. O. Segerdahl*. Operation analysis: experience and concept: *R. A. Porter*. Research in collective risk theory: *G. Arfwedson*.

SWITZERLAND—

Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker—

Vol. 55, No. 3—Statistische Schätzungen mit Quantilen: *A. Dalcher*. Über zusammengesetzte Poisson-Prozesse und ihre Anwendungen in der Unfallversicherung: *M. Hofmann*. Zur mathematischen Theorie der Einkommensverteilung in Abhängigkeit von Alter und Zeit: *F. Leuenberger*.

Schweizerische Zeitschrift für Volkswirtschaft und Statistik—

September—Contribution au calcul simplifié des pertes de potentiel—vie causées par l'avancement en âge de la population: *R. Sailer*. Probleme des Lebenskostenindex in Grossbritannien: *W. Zeller*.

INTERNATIONAL—

International Labour Review—

October 1955—Pensionable age under Old-Age Pension schemes: *D. S. Gerig*.
November 1955—Productivity and economic development in Latin America: *J. Franco*.
December 1955—Automation—some social aspects: *H. de Bivort*.

World Health Organization Bulletin—

Vol. 13, No. 3—The use and results of diptheria immunization: *L. Greenberg*. Large-scale field trials of active immunizing agents, with special reference to vaccination against pertussis: *W. C. Cockburn*.

Vol. 13, No. 4—L'absentéisme dans l'industrie: *S. Forssman*. Sickness absenteeism: *G. J. Fortuin*. The employment of elderly persons: *T. A. Lloyd-Davies*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

LIST OF ADDITIONS TO THE LIBRARY

Since the issue of Part IV, 1955, the Society has received the publications enumerated below:

I.—OFFICIAL PUBLICATIONS.

United Kingdom. *Agricultural Research Council.* Quantitative inheritance: papers read at a colloquium held at the Institute of Animal Genetics, Edinburgh University, under the auspices of the Agricultural Research Council. April 4th to 6th, 1950; edited by E. C. R. Reeve and C. H. Waddington. London, H.M.S.O., 1952. 151p. 9¾″. 20*s.*

Atomic Energy Authority. First annual report 1954–55. London, H.M.S.O., 1955. vii, 35p. 9¾″. 2*s.*

Colonial Office. An economic survey of the colonial territories. Vol. V. The Far Eastern territories: the Federation of Malaya, Singapore, Hong Kong, Brunei, North Borneo and Sarawak. London, H.M.S.O., 1955. 159p. 3 fold. maps. 13½″. 27*s.* 6*d.* (Colonial No. 281–5)

Commonwealth Economic Committee. Jute manufactures: a memorandum on production, prices and trade. London, H.M.S.O., 1955. iv, 42p. 9¾″. 2*s.* 6*d.*

Commonwealth Relations Office & Colonial Office. High Commission territories: economic development and social services. London, H.M.S.O., 1955. 19p. map. 9¾″. 9*d.* (Cmd. 9580)

Department of Scientific and Industrial Research. Scientific research in British universities 1954–1955. London, H.M.S.O., 1955. v, 572p. 8″. 12*s.* 6*d.*

General Register Office. Census 1951. England and Wales. Index of place names. London, H.M.S.O., 1955. 2 v. 12¾″. 63*s.*

Medical Research Council. The diagnosis and treatment of haemophilia and its related conditions; R. G. MacFarlance and Rosemary Biggs. London, H.M.S.O., 1955. viii, 23p. 9½″. 2*s.* 6*d.* (M.R.C. Memorandum No.32)

Lung function in coalworkers' pneumoconiosis; J. C. Gilson and P. Hugh-Jones. London, H.M.S.O., 1955. xii, 266p. 15 plates. 9¾″. 21*s.* (M.R.C. Spec. Rep. Ser. 290)

Ministry of Education. A short history of the Science Museum; compiled by F. Greenaway. London, H.M.S.O., 1951. 18p. 4 plates. 9¾″. 1*s.*

Ministry of Housing and Local Government. Report of the Ministry of Housing and Local Government for the period 1950/51–1954. London, H.M.S.O., 1955. xii, 227p. 9¾″. 7*s.* (Cmd. 9559)

Ministry of Labour and National Service. National Advisory Committee on the Employment of Older Men and Women. Second report. London, H.M.S.O., 1955. 28p. 9¾″. 1*s.* 9*d.* (Cmd. 9628)

Monopolies and Restrictive Practices Commission. Report on the supply and export of certain semi-manufactures of copper and copper-based alloys. London, H.M.S.O., 1955. viii, 232p. 9½″. 7*s.* 6*d.*

Report on the supply and export of pneumatic tyres. London, H.M.S.O., 1955. v, 233p. 9¾″. 6*s.* 6*d.*

Office of the Lord President of the Council. Advisory Council on Scientific Policy. Committee on Scientific Manpower. Report on the recruitment of scientists and engineers by the engineering industry. London, H.M.S.O., 1955. 31p. 9¾″. 1*s.* 3*d.*

Post Office. Report on Post Office development and finance. London, H.M.S.O., 1955. 19p. 9¾″. 9*d.* (Cmd. 9576)

Royal Commission on the Civil Service. Report, 1953–55. London, H.M.S.O., 1955. viii, 238p. 9¾″. 6*s.* 6*d.* (Cmd. 9613)

Scotland. General Registry Office. Life tables 1950–52: supplement to the 99th annual report of the Registrar General for Scotland. Edinburgh, H.M.S.O., 1954. 24p. 9½″. 1*s.* 3*d.*

California. *Department of Fish and Game. Marine Research Committee.* California cooperative oceanic fisheries investigations: progress report, 1 July, 1953 to 31 March, 1955. [Sacramento, 1955.] 52p. 11″.

Council of Europe. *Consultative Assembly.* Focus on East–West relations: a policy for Europe. Strasbourg, 1956. 252p. 10½″.

Secretariat-General. Statistical data, for the use of the Council of Europe. Vol. 2. Industrial statistics. Strasbourg, 1955. 87p. 9¼″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cuba. *Consejo Nacional de Economia.* El programa economico de Cuba. La Habana, 1955. 84p. 8¼″.

Denmark. *Arbejds- og Socialministeriet.* Personer under Østifternes Åndssvageanstalts forsorg: en redegørelse ved arbejds- og social-ministeriernes statsvidenskabelige konsulent. Persons under the care of Østifternes Åndssvageanstalt. English summary. København, 1955. 50p. 9″. (Økonomisk-statistiske Undersøgelsen Nr. 20)

Statistiske Departement. Folketaelling 1950. Aldersfordelingen i 1950 og i de kommende aar ... Age distribution in 1950 and the following decades: sex, age and marital status of the 1950 census. København, 1955. 155p. 9¼″. (Statistiske Meddelelser 4. 162. 1)

Nationalindkomsten 1938 og 1947–54. National accounting 1938 und 1947–54. København, 1955. 126, 75p. 3 fold. tables. 9¼″. *Kr.* 3.50. (Statistiske Meddelelser 4. 160. 2.)

Denmark, Finland, Iceland, Norway & Sweden. *Ministers of Social Affairs.* Samordning af de nordiske landes statistik vedrørende den sociale lovgivning. Co-ordination of social welfare statistics in the northern countries. København, Jørgensen, 1955. 136p. 9½″. (Statistical Reports of the Northern Countries 2). Summary in English.

Eire. *Central Statistics Office.* Census of population of Ireland 1951. Vol. III. Part II. Industries and industrial status. Dublin, Stationery Office, 1954. viii, 213p. 12¾″. 7*s.*

France. *Institut National de la Statistique.* Annuaire statistique de la Guadeloupe 1949–1953. Paris, 1954. 132p. fold. map. 10½″.

Frankfurt am Main. *Statistisches Amt.* Statistisches Jahrbuch, Frankfurt am Main, 1954. Frankfurt, 1955. xiv, 186p. 10¼″.

Germany. *Statistisches Bundesamt.* Einführung in die methodischen und systematischen Grundlagen der Volks- und Berufszählung vom 13.9.50. Stuttgart-Köln, 1955. 110p. 11¾″. DM5.50. (Statistik der Bundesrepublik Deutschland Band 34.)

Gebäude- und Wohnungszählung in der Bundesrepublik Deutschland vom 13.9.1950. Heft 2. Hauptergebnisse nach Kreisen. Stuttgart-Köln, 1955. 82p. 11¾″. DM4.20. (S.B.D. Band 38.)

Die Kostenstruktur in der gewerblichen Wirtschaft und in ausgewählten freien Berufen (Ergebnisse der Kostenstrukturerhebung 1950). Heft 2. Handwerk. 95p. DM4.60. Heft 4. Einzelhandel. 119p. DM6. Heft 6. Verkehrsgewerbe. 28p. DM1. Stuttgart-Köln, 1955. 11¾″. (S.B.D. Band 49, Heft 2, 4, 6.)

Wirtschaftskunde der Bundesrepublik Deutschland. Wiesbaden, 1955. 561p. maps (1 fold.) illus. 8¼″.

India. *Office of the Registrar General.* Census of India, 1951. Volume 1. India. Part II-A—Demographic tables. Delhi, 1955. [iv], 412p. 11″. 5*s.*

Census of India. Paper No. 1. 1955. Sample census of births and deaths—1953–54; Uttar Pradesh. vi, 185p. fold. map. 11¼″. 3*s.* Paper No. 2. 1955. Sample census of births and deaths—1952–53; India. ii, xii, 540p. 11″. 3*s.*

Italy. *Istituto Centrale di Statistica.* Classificazioni delle malattie e cause di morte. Roma, 1955. 225p. 8½″.

Karinthia. *Amt der Kärntner Landesregierung.* Statistisches Handbuch des Landes Kärnten. 1. Jahrgang. Klagenfurt, 1955. xiii, 104p. 11¼″.

Kenya. *East African Statistical Department.* Kenya agricultural census, 1954 (Highlands and Asian settled area). [Nairobi], 1955. 30p. 12¾″. The first annual agricultural census since 1940.

Mozambique. *Repartição Tecnica de Estatística.* Recenseamento agrícola 1951. Lourenço Marques, 1955. 785p. map. 10½″.

Norway. *Statistisk Sentralbyrå.* Nasjonalregnskap: teoretiske prinsipper. National accounts: theoretical principles; by Odd Aukrust. Oslo, 1955. 123p. 9½″. (Samfunnsøkonomiske Studier 4). English summary.

Oporto. *Durius.* O Porto em 1950: subsidios demograficos, economicos e financeiros referentes a cidade do Porto; organizados por Durius. Porto, 1955. 44p. fold. diagr. illus. 9¾″.

Organisation for European Economic Co-operation. Trend of the labour market in member countries in recent years: report by the Manpower Committee of O.E.E.C. Paris, 1955. 53p. 10¾″. Typescript.

European Productivity Agency. Production measurement. I. Concepts. Paris, O.E.E.C., 1955. 143p. 9½″. 7*s.* (Project No. 235)

Portugal. *Instituto Nacional de Estatïstica.* Indices de salarios por profisseõs para a cidade de Lisboa. Lisboa, 1955. 53p. 10¼″. (Estudos No. 28)

Rome. *Pio Istituto di Santo Spirito ed Ospedali Riuniti di Roma.* La popolazione ospedaliera (sesso—età—cause morbose) negli anni 1934–1950. Roma, Istituto Poligrafico dello Stato, 1954. 222p. 10½″. Summary in English.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Sweden. *Kommerskollegium.* [Census of production, Distribution and Services.] 1951 års företagsräkning. Stockholm, Kommerskollegium, 1955. 386p. 9½″. *Kr.* 6. Summary in English.

Statistiska Centralbyrån. Jordbrukarnas tillgangar och skulder, den 31 december 1952. Stockholm, 1955. x, 31°, 39p. 9½″.

Switzerland. *Eidgenössisches Statistisches Amt.* Eidgenössische Volkszählung 1. Dezember 1950. Band 10. Kanton Luzern. 193p. Band 26. Schweizerbürger nach Heimatkantonen und -gemeinden. 81p. Bern, 1955. 11¾″. (Statistische Quellenwerke der Schweiz, Heft 279, 280)

Vienna. *Statistisches Amt.* Die Wiener in der Konjunktur. Wien, Magistrat der Stadt, 1955. 17p. 11¾″. (Mitteilungen aus Statistik und Verwaltung der Stadt Wien, 1955, Sonderheft 4.)

United Nations. *Department of Economic and Social Affairs.* Economic developments in the Middle East, 1945 to 1954: supplement to *World Economic Report, 1953–54.* New York, 1955. x, 236p. 10¾″. 17*s.* 6*d.*

Economic survey of Latin America 1954; prepared by the secretariat of the Economic Commission for Latin America. New York, 1955. xvi, 203p. 11″. 17*s.* 6*d.*

Review of economic activity in Africa 1950 to 1954; supplement to *World Economic Report, 1953–54.* New York, 1955. xii, 146p. 10¾″. 11*s.*

Economic Commission for Europe & Food and Agriculture Organization. Prices of agricultural products and fertilizers 1954/55. Geneva, U.N., [1955]. [vii], 50, 2, 18p. 11″. 4*s.* 6*d.* The fifth of a series of annual studies on the same subject.

Food and Agriculture Organization. Food balance sheets. Second issue. Rome, F.A.O., 1955. 113 l. 8½″ × 10¾″. Loose leaf.

Problems of food and agricultural expansion in the Far East. Rome, 1955. vii, iii, 78p. 9″. 3*s.* 9*d.*

General Agreement on Tariffs and Trade. Further tentative estimates of the commodity composition of international trade and presentation of a world network of trade in constant prices. Geneva, 1955. 17 l. 4 fold. tables on 7 leaves. 11″. (G.A.T.T. Trade Intelligence Paper No. 5)

Statistical Office. Handbook of vital statistics methods. New York, 1955. iv, 258p. 9 fold. tables in pocket. 11″. 17*s.* 6*d.* (Statistical Papers Series F, No. 7)

United States. *Bureau of Labor Statistics.* Science and engineering in American industry: preliminary report on a survey of research and development costs and personnel in 1953–1954; prepared for the National Science Foundation by the U.S. Department of Labor, Bureau of Labor Statistics. Washington, Superintendent of Documents, 1955. v, 41p. 10¼″. 30*c.*

National Bureau of Standards. Table of hyperbolic sines and cosines x = 2 to x = 10. Washington, Superintendent of Documents, 1955. v, 81p. 10¼″. 55*c.* (Appl. Math. Ser. 45)

Table of the descending exponential x = 2.5 to x = 10. Washington, Superintendent of Documents, 1955. v, 76p. 10¼″. 50*c.* (Appl. Math. Ser. 46). A supplement to *Tables of the exponential function* e^x; prepared by the Works Project Administration, New York, 1939.

Office of Business Economics. National income, 1954 ed: a supplement to the *Survey of Current Business*; prepared by the National Income Division. Washington, Superintendent of Documents, 1954. vi, 249p. illus. 11½″. $1.50.

West Bengal. *State Statistical Bureau.* Comprehensive statistics of staff employed under the government of West Bengal (as on 28th February, 1954). Alipore, 1955. ii, 211p. 13¼″. For offical use only.

Report on the sample surveys for estimating the acreages and yield rates of rabi crops in West Bengal 1953–54. Vol. III. Part III. Alipore, 1955. [ii], 53p. 9¾″. 1*s.* 8*d.*

Yugoslavia. *Federal Statistical Office.* Konacni rezultati popisa stanovnistva od 15 marta 1948 godine. Knjiga V. Stanovnistvo po pismenosti. Final results of the population census of March 15th, 1948. Volume V. Population by literacy. Beograd, 1955. lxiv, 384p. 11½″.

II.—AUTHORS AND MISCELLANEOUS.

Abel-Smith, Brian & Titmuss, Richard M. The cost of the National Health Service in England and Wales. Cambridge, University Press, 1956. xx, 176p. 9½″. 27*s.* 6*d.* (N.I.E.S.R. Occ. Pap. 18)

Allais, Maurice. Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine. [Paris], Imprimerie Nationale, [1955]. 55p. 10½″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Almond, Joyce. A note on the χ^2 test applied to epidemic chains. *Biometrics* (1954), **10**, 459–477. $9\frac{1}{4}''$.

American Public Health Association. On the use of sampling in the field of public health; prepared by the Committee on Sampling Techniques in Public Health Statistics, Statistics Section. *Amer. J. Publ. Hlth.* (1954), **44**, 719–740. $9\frac{3}{4}''$.

Association of Incorporated Statisticians. Modern sample survey methods: papers read at a week-end school held in December, 1953. London, 1955. ii, 117p. $8\frac{1}{2}''$. 10*s.* 6*d.* Supplement to *The Incorporated Statistician*, **5**, no. 4, April, 1955.

Barnett, H. A. R. The components of mortality. *J. Inst. Actu.* (1955), **81**, 105–149. fold. tab. $8\frac{3}{4}''$.

The variance of the product of two independent variables and its application to an investigation based on sample data. *J. Inst. Actu.* (1955), **81**, 190. $8\frac{3}{4}''$.

Barr, A. Employment of recovered tuberculous patients. *Brit. J. Prev. Soc. Med.* (1955), **9**, 101–103. $9\frac{3}{4}''$.

Bartlett, M. S. Approximate confidence intervals. III. A bias correction. *Biometrika* (1955), **42**, 201–204. $10\frac{3}{4}''$.

The statistical analysis of stochastic processes. *Colloq. sur l'Analyse Statist., Bruxelles*, 1954, 113–132.

Bartlett, M. S. & Medhi, J. On the efficiency of procedures for smoothing periodograms from time series with continuous spectra. *Biometrika* (1955), **42**, 143–150. $11\frac{3}{4}''$.

Behrens, C. B. A. Merchant shipping and the demands of war. London, H.M.S.O., & Longmans, Green, 1955. xix, [ii], 494p. photos, fold. maps, fold. diagrs. $9\frac{1}{2}''$. 35*s.* (History of the Second World War, U.K. Civil Series)

Blumen, I., Kogan, Marvin & McCarthy, Philip J. The industrial mobility of labor as a probability process. Ithaca, N.Y., Cornell University, 1955. xii, 163p. 9″. (Cornell Studies in Industrial and Labor Relations, Vol. VI)

Boag, J. W. The presentation and analysis of the results of radio-therapy. Part I.—Part II. Mathematical theory. *Brit. J. Radiol. N.S.* (1948), **21**, 128–138, 189–203. $9\frac{3}{4}''$.

Bochner, Salomon. Harmonic analysis and the theory of probability. Berkeley & Los Angeles, University of California Press, 1955. London, Cambridge University Press. viii, 176p. $8\frac{1}{2}''$. 35*s.*

Braithwaite, R. B. Theory of games as a tool for the moral philosopher. Cambridge, University Press, 1955. 76p. $7\frac{1}{4}''$. 6*s.* Inaugural lecture by the Knightbridge Professor of Moral Philosophy.

Brooks, Samuel H. The estimation of an optimum subsampling number. *J. Amer. Statist. Ass.* (1955), **50**, 398–415. $9\frac{1}{4}''$.

Bush, Robert R. & Mosteller, Frederick. Stochastic models for learning. New York, Wiley, 1955. London, Chapman & Hall. xvi, 365p. 9″. 72*s.*

Butler, D. E. The British general election of 1955. London, Macmillan, 1955. [viii], 236p. illus. $8\frac{1}{2}''$. 24*s.*

Cambridge. University. *Department of Applied Economics.* Measuring social mobility; S. J. Prais. *J.R. Statist. Soc. A.* (1955), **118**, 56–66. $9\frac{1}{2}''$. (Repr. Ser. 97)

Some estimation problems in demand analysis; Richard Stone, J. Aitchison & J. A. C. Brown. *Incorporated Statistician* (1955), **5**, 165–177. 13p. $8\frac{1}{2}''$. (Repr. Ser. 98)

An econometric study of supply and demand for New Zealand's exports; by A. R. Bergstrom. *Econometrica* (1955), **23**, 258–276. 10″. (Repr. Ser. 99)

Transaction models with an example based on the British national accounts; by Richard Stone. *Accounting Res.* (1955), **6**, 202–226. 24p. $9\frac{3}{4}''$. (Repr. Ser. 100)

The formal theory of social mobility; by S. J. Prais. *Popul. Stud.* (1955), **9**, 72–81. 10″. (Repr. Ser. 101)

Chapman, Dennis. The home and social status. London, Routledge & Kegan Paul, 1955. xvii, 301p. $8\frac{1}{2}''$.

Clarke, F. Le Gros & Dunne, Agnes C. Ageing in industry: an inquiry, based on figures derived from census reports, into the problem of ageing under the conditions of modern industry. London, Nuffield Foundation, 1955. xii, 146, [viii]p. $8\frac{1}{2}''$. 6*s.*

Clarke, Geoffrey M. A design for testing several treatments under controlled environmental conditions. *Appl. Statist.* (1955), **4**, 199–206. 10″.

Clarke, Geoffrey M. & Corke, A. T. K. Sources of variability in sampling blackcurrants for leaf spot. *Long Ashton Research Station Ann. Rep. 1954* (1955), 199–202. $9\frac{1}{2}''$.

Clarke, L. Harwood. A note book in applied mathematics. London, Heinemann, 1956. x, 163p. $7\frac{3}{4}''$. 9*s.* 6*d.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cochran, William G. Analysis of records with a view to their evaluation. *Family Hlth. Maintenance Demonstration* (1953), 228–236. 9p. 9″.

The combination of estimates from different experiments. *Biometrics* (1954), **10**, 101–129. 9¼″.

The present structure of the Association. *J. Amer. Statist. Ass.* (1954), **49**, 1–12. 9¼″. Presidential Address.

The relation between simple transforms and maximum likelihood solutions. *Biometrics* (1954), **10**, 144–147. 9¼″.

Research techniques in the study of human beings. *Milbank Mem. Fd. Quart. Publ.* (1955), **33**, 121–136. 9″.

Some methods for strengthening the common tests. *Biometrics* (1954), **10**, 417–451. 9¼″.

A test of a linear function of the deviations between observed and expected numbers. *J. Amer. Statist. Ass.* (1955), **50**, 377–397. 9¼″.

Cox, D. R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proc. Camb. Phil. Soc.* (1955), **51**, 433–441. 10″.

Croxton, Frederick E. Elementary statistics with applications in medicine. New York, Prentice-Hall, 1953. London, Bailey Bros. & Swinfen. vii, 376p. 8½″. 80*s*.

Das, S. C. The fitting of truncated type III curves to daily rainfall data. *Aust. J. Phys.* (1955), **8**, 298–304. 9½″.

Davidson, Ralph Kirby. Price discrimination in selling gas and electricity. Baltimore, Johns Hopkins Press, 1955. 254p. 9″.

Davis, David E. & Zippin, Calvin. Planning wildlife experiments involving percentages. *J. Wildlife Mgmt.* (1954), **18**, 170–178. 9¾″.

Davis, Harold T. Political statistics . . . Evanston, The Principia Press of Illinois, 1954. xv, [iii], 364p. 8½″. 42*s*.

Davis, Hiram S. Productivity accounting. Philadelphia, University of Pennsylvania Press, 1955. xii, 194p. 9″. 40*s*. (Wharton School of Finance and Commerce Res. Stud. 37)

Dobson, R. Warwick. An introduction to cost accountancy. Vol. III. London, Gee & Co., 1954. 352p. fold. diagrs. 8½″. 37*s*. 6*d*.

Dornbusch, Sanford M. & Schmid, Calvin F. A primer of social statistics. New York & London, McGraw-Hill, 1955. ix, 251p. illus. 9″. 35*s*. 6*d*.

Dudding, B. P. Drafting specifications based on limiting the number of defectives permitted in small samples. British Standards Institution, 1955. 10p. fold. diagr. 8½″. 2*s*. 6*d*. (B.S. 2635: 1955)

Dudding, B. P. & Jennett, W. J. British Standard control chart technique when manufacturing to a specification: with special reference to articles machined to dimensional tolerances. British Standards Institution, 1955. 77p. 8½″. 10*s*. 6*d*. (B.S. 2564: 1955)

Duncan, David C. A new method of recording labour losses. *The Manager* (1955), January. [8p.] 11¾″.

Federation of British Industries. A digest of five offical reports on the taxation of profits and income: representing six years' examination of the law by the Royal Commission and the two Millard Tucker Committees; [compiled by Frank Bower]. London, F.B.I., 1955. vii, 92p. 8½″. 5*s*.

Fieller, E. C., Lewis, T., & Pearson, E. S. Correlated random normal deviates: 3,000 sets of deviates, each giving 9 random pairs with correlations 0·1 (0·1) 0·9, compiled from Herman Wold's table of random normal deviates (Tract No. XXV). Cambridge, University Press, 1955. xv, 60p. 9¼″. 10*s*. 6*d*. (Tracts for Computers No. XXVI)

Finney, D. J. Experimental design and its statistical basis. London, Cambridge University Press, 1955. xi, 168p. 8½″. 30*s*.

Finney, D. J., Hazlewood, T. & Smith, Marian J. Logarithms to base 2. *J. Gen. Microbiol.* (1955), **12**, 222–225. 10¼″.

Fisher, *Sir* Ronald A. The design of experiments. 6th ed. Edinburgh & London, Oliver & Boyd, 1953. xv, 244p. 8¾″. 12*s*. 6*d*.

Fix, Evelyn & Hodges, J. L., *Jr*. Significance probabilities of the Wilcoxon test. *Ann. Math. Statist.* (1955), **26**, 301–312. 10″.

Forte, Francesco. Note sulla natura economica e sugli effetti delle imposte sul valore aggiunto. Milano, Unione Tipografica, 1955. 30p. 9½″. (Inst. Sci. econ. statist. Quad. 21)

Frisch, Ragnar. Principles of linear programming: with particular reference to the double gradient form of the logarithmic potential method. Oslo, Universitetets Socialøkonomiske Institutt, 1954. v, 219p. 11¾″.

Gabor, André. Concept of statistical freedom and its application to social mobility. *Popul. Stud.* (1955), **9**, 82–95. 10″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Gani, J. & Moran, P. A. P. The solution of dam equations by Monte Carlo methods. *Aust. J. Appl. Sci.* (1955), **6**, 267–273. 9½″.

Guinness, Arthur, *Son & Co., Ltd.* The Guinness book of records. London, Guinness Superlatives Ltd., 1955. x, 198p. 17 plates. illus. 9¾″.

Gunzert, Rudolf. Regionalstatistik und repräsentative Methode: eine Schicksalsfrage der deutschen Städtestatistik. Frankfurt am Main, *Statistische Monatsberichte* (1955), Heft 10. 18p. 11¾″.

Haberler, Gottfried. A survey of international trade theory. (A translation from the German, by W. Michael Blumenthal, of the article "Aussenhandel [Theorie]," in Handwörterbuch der Sozialwissenschaften, 1954. Revised and enlarged by the author.) Princeton, University, International Finance Section, 1955. [vi], 68p. 10″. (Special Papers in International Economics 1)

Hannan, E. J. A test for singularities in Sydney rainfall. *Aust. J. Phys.* (1955), **8**, 289–297. 9½″.

Hanson, Kermit O. Managerial statistics. New York, Prentice-Hall, 1955. xiv, 306p. illus. 8½″. 51*s*.

Harvard University. *Computation Laboratory.* Tables of the cumulative binomial probability distribution. Cambridge, Mass., Harvard University Press, 1955. London, Cumberlege. lxi, 503p. 10½″. 63*s*. (Annals of the Computation Laboratory 35)

Haycocks, H. W. & Perks, W. Mortality and other investigations. Cambridge, University Press, for the Institute of Actuaries and the Faculty of Actuaries, 1955. ix, 164p. 8½″. 20*s*.

Hodges, J. L., *Jr.* A bivariate sign test. *Ann. Math. Statist.* (1955), **26**, 523–527. 10″.

Galton's rank-order test. *Biometrika* (1955), **42**, 261–262. 10¾″.

Hoffmann, Walther G. British industry 1700–1950. Translated by W. O. Henderson & W. H. Chaloner. Oxford, Blackwell, 1955. xxiii, 338p. fold. tables. Originally published as *Wachstum und Wachstumsformen der englischen Industriewirtschaft.* Jena, 1940. (Probleme der Weltwirtschaft, vol. 63). The author has incorporated some new material.

Hogben, Lancelot. Man must measure: the wonderful world of mathematics. London, Rathbone Books, 1955. [3] 7–69p. illus. 12¾″. 15*s*.

Hooper, Shadrach G. The finance of farming in Great Britain. London, Europa Publications, [1956]. 247p. map. 8½″. 25*s*.

Institute of Bankers. International banking and foreign trade; being lectures delivered at the eighth International Banking Summer School, Christ Church, Oxford, September 1955. London, Europa Publications, for the Institute of Bankers, [1955]. 226p. 8½″. 20*s*.

International Federation of Cotton and Allied Textile Industries. The improvement of Ashmouni cotton: official report of the Ashmouni Study Mission to Egypt, 1955. Manchester, I.F.C.A.T.I., 1955. v, 41p. 9½″.

Ireson, William Grant & Grant, Eugene L., *editors.* Handbook of industrial engineering and management. Englewood Cliffs, N. J., Prentice-Hall, 1955. London, Bailey Bros. & Swinfen. vii, [3], 1203p. 9″. 128*s*.

Johnson, L. G. Richard Jones reconsidered: a centenary tribute. Doncaster, printed for private circulation, 1955. 14p. 8¾″.

Joint Committee of the Medical Research Council & Nuffield Foundation on Clinical Trial of Cortisone, A.C.T.H. & Other Therapeutic Measures in Chronic Rheumatic Diseases. A comparison of cortisone and aspirin in the treatment of early cases of rheumatoid arthritis: a second report. *Brit. Med. J.* (1955), **2**, p.695. 15p. 8½″.

Jowett, G. H. The comparison of means of industrial time series. *Appl. Statist.* (1955), **4**, 32–46. 10″.

Karpinos, Bernard P. Evaluation of the physical fitness of present-day inductees. *U.S. Forces Med. J.* (1953), **4**, 415–430. 9″.

Karpinos, Bernard D. & Grossman, Harold A. Prevalence of left-handedness among selective service registrants. *Hum. Biol.* (1953), **25**, 36–49. 8¾″.

Klatzmann, Joseph. La localisation des cultures et des productions animales en France. Paris, Imprimerie Nationale, 1955. 477p. 1 fold. map and 2 loose maps. 10″.

Kopal, Zdeněk. Numerical analysis; with emphasis on the application of numerical techniques to problems of infinitesimal calculus in single variables. London, Chapman & Hall, 1955. xiv, 556p. 8¾″. 63*s*.

Kuczynski, Jürgen. Die Geschichte der Lage der Arbeiter in den Vereinigten Staaten von Amerika von 1775 bis in die Gegenwart. 1. Teil. Von der Befreiung vom kolonialen Joch bis zur Zeit vor der Herrschaft des Monopolkapitals, 1775 bis 1897. 247p. DM. 6.30. 2. Teil. Die Herrschaft des Monopolkapitals seit 1898. 291p. DM7. 3. verbesserte Aufl. Berlin,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Tribüne, 1955. 8¼″. (Die Geschichte der Lage der Arbeiter unter dem Kapitalismus Band 3, Teil 1 & 2)

Die Geschichte der Lage der Arbeiter in Deutschland von 1789 bis in die Gegenwart. 2. Teil. Westdeutschland seit 1945. 3. verbesserte Aufl. Berlin, Tribüne, 1955. 378p. 8¼″. DM 7.50. (Band 2, Teil 2)

Die Geschichte der Lage der Arbeiter in Frankreich von 1789 bis in die Gegenwart. 1. Teil. 1789 bis 1830. 287p. DM6.85. 2. Teil. Seit 1830. 348p. DM7.50. 2. verbesserte Aufl. Berlin, Tribüne, 1955. 8¼″. (Band 6, Teil 1 & 2)

Die Geschichte der Lage der Arbeiter im englischen Kolonialreich. 2. verbesserte Aufl. Berlin, Tribüne, 1955. 294p. 8¼″. DM7. (Band 5)

Die Theorie der Lage der Arbeiter. 3. verbesserte Aufl. Berlin, Tribüne, 1955. 235p. 8¼″. (Band 7)

Kulldorff, Gunnar. Migration probabilities. Lund, Gleerup, 1955. 46p. 9¼″. *Kr.* 5. (Lund Studies in Geography B. 14)

Lasheras-Sanz, Antonio. Aspectos economicos y sociales del seguro. La Habana, Consejo Nacional de Economia, 1955. 150p. 8¼″. (Estudios e Investigaciones Economicas 13)

Lehmann, E.L. Ordered families of distributions. *Ann. Math. Statist.* (1955), **26**, 399–419. 10″.

Lehmann, E.L. & Scheffé, Henry. Completeness, similar regions, and unbiased estimation—part II. Sankhyā (1955), **15**, 219–236. 11¼″.

Lewis, W. Arthur. The theory of economic growth. London, Allen & Unwin, 1955. 453p. 8½″. 30*s.*

Loève, Michel. Calcul des probabilités.—Relations entre lois limites. *C.R. Acad. Sci.* (1954). 3p. 11¾″.

Calcul des probabilités.—Termes variationnels dans le modèle limite ventral. *C.R. Acad. Sci.* (1955). 3p. 10¾″.

Probability theory foundations, random sequences. New York, Van Nostrand, 1955. London, Macmillan. xv, [ii], 515p. 9″. 90*s.*

London. University. *University College. Department of Statistics.* Elementary statistical exercises. Part I; [prepared by F. N. David]. London, The Department, 1952. [iv], 91p. 11″. 6*s.* 6*d.*

Ma, Ronald A. & Szczepanik, Edward F. The national income of Hong Kong 1947–1950. Hong Kong, University Press, 1955. London, Cumberlege. x, 69p. 8½″. 12*s.* 6*d.*

McVittie, George C. Relativity and the statistical theory of the distribution of galaxies. *Astr. J.* (1955), **60**, 105–115. 10½″.

Marriott, F. H. C. & Pope, J. A. Bias in the estimation of autocorrelations. *Biometrika* (1954), **41**, 390–402. 10¾″.

Martin, W. J. Vital statistics of the county of London in the years 1901 to 1951. *Brit. J. Prev. Soc. Med.* (1955), **9**, 126–134. 9¾″.

Marx, Norbert. Etude sur le prix de revient de deux coefficients opératoires (K 50 et K 100) dans divers établissements hospitaliers. Paris, Caisse Régionale de Sécurité Sociale, [1952]. [1], ii, 52 l. 25 diagrs., 3 folding. 10¾″.

Etude sur l'évolution des dépenses annuelles du risque "maladie" à la Caisse Centrale de Sécurité Sociale de la Région Parisienne de 1932 à 1951; par le docteur Norbert Marx . . . et Essai d'interprétation des données comptables; par Monsieur Marc . . . Paris, Caisse Régionale de Sécurité Sociale, [1952]. [1], ii, 26 l. 59 diagrs., 39 folding. 10¾″.

Le risque "longue maladie": représentation graphique de son évolution pendant les années 1947 à 1951 à la Caisse Centrale de Sécurité Sociale de la Région Parisienne. Paris, Caisse Régionale de Sécurité Sociale, [1952]. [1], ii, 31 l. 24 diagrs., 14 folding. 10¾″.

Meehl, Paul E. Clinical versus statistical prediction: a theoretical analysis and a review of evidence. Minneapolis, University of Minnesota Press, 1954. x, 149p. 8½″. 24*s.*

Meier, Paul. Analysis of simple lattice designs with unequal sets of replications. *J. Amer. Statist. Ass.* (1954), **49**, 786–813. 9¼″.

Note on estimation in a Markov process with constant transition rates. *Hum. Biol.* (1955), **27**, 121–124. 9¼″.

Meier, Paul & Zierler, Kenneth L. On the theory of the indicator-dilution method for measurement of blood flow and volume. *J. Appl. Physiol.* (1954), **6**, 731–744. 10″.

Merrell, Margaret & Shulman, Lawrence E. Determination of prognosis in chronic disease, illustrated by systematic lupus erythematosus. *J. Chronic Diseases* (1955), **1**, 12–32. 10″.

Miani-Calabrese, Donato. Un metodo di costruzione di un indice dello sviluppo economico. *Riv. Polit. econ.* (1955), **15**, III Serie, Fasc. 7–8. 13p. 9¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Milbank Memorial Fund. Catalogue of publications 1905–1955. New York, 1955. 65p. 9″.

Miller, Herman P. Income of the American people. New York, Wiley, for the Social Science Research Council in cooperation with the U.S. Bureau of the Census, 1955. London, Chapman & Hall. xvi, 206p. 9″. 44*s*. (Census Monograph Series)

Moran, P. A. P. A probability theory of dams and storage systems. *Aust. J. Appl. Sci.* (1954), **5**, 116–124. 9½″.

A probability theory of dams and storage systems: modifications of the release rules. *Aust. J. Appl. Sci.* (1955), **6**, 117–130. 9½″.

Morgenstern, Oskar. The validity of international gold movement statistics. Princeton, University, International Finance Section, 1955. [vi], 42p. fold. diagr. 9″. (Special Papers in International Economics 2)

National Institute of Economic and Social Research. Edith; by Brian Tew. *Three Banks Rev.* (1955), no. 26, 27–43. 17p. 8½″. (Repr. Ser. 3). For private circulation.

The I.C.F.C. revisited; by Brian Tew. Economica (1955), **22**, 218–233. 16p. 9¾″. (Repr. Ser. 4). For private circulation.

Neale, E. P. Guide to New Zealand official statistics. (3rd ed.—enlarged and almost entirely re-written.) Christchurch, [etc.], London, Whitcombe & Tombs, 1955. 195p. 8½″. 16*s*. (Auckland University College Texts 2)

Netherlands. *Centraal Planbureau.* De motorisatie van de Nederlandse landbouw; [door] A. W. G. Koppejan. *Landbouwk. Tijdschr.* (1955), **67**, 315–324. 9½″. (Overdrukken 40). Summary in English.

Recent experiences with the Munich business test: an expository article; by H. Theil. *Econometrica* (1955), **23**, 184–192. 9½″. (Overdrukken 41)

De huidige sociale zekerheid van een werknemer in Nederland; door Dr. J. G. M. Delfgaauw & M. de Willigen. *Social Maandblad Arbeid* (1955), **10**, 552–564. 13p. 9″. (Overdrukken 42)

Een verkenning der economische toekomstmogelijkheden van Nederland 1950–1970. 's-Gravenhage, 1955. 144p. 9½″.

North Carolina. *Consolidated University. Institute of Statistics.* Doubly balanced incomplete block designs for experiments in which the treatment effects are correlated; by Lyle D. Calvin. *Biometrics* (1954), **10**, 61–88. 9″. (Repr. Ser. 49)

On multivariate distribution theory; by I. Olkin & S. N. Roy. *Ann. Math. Statist.* (1954), **25**, 329–339. 9¾″. (Repr. Ser. 51)

A confidence region for the solution of a set of simultaneous equations with an application to experimental design; by G. E. P. Box & J. S. Hunter. *Biometrika* (1954), **41**, 190–199. 10½″. (Repr. Ser. 52)

A useful theorem in matrix theory; by S. N. Roy. *Proc. Amer. Math. Soc.* (1954), **5**, 635–638. 9¼″. (Repr. Ser. 53)

A comparison of stratified two-stage sampling systems; by A. R. Sen, R. L. Anderson & A. L. Finkner. *J. Amer. Statist. Ass.* (1954), **49**, 539–558. 9″. (Repr. Ser. 54)

On the distribution of the ratio of the *i*th observation in an ordered sample from a normal population to an independent estimate of the standard deviation; by K. C. S. Pillai & K. V. Ramachandran. *Ann. Math. Statist.* (1954), **25**, 565–572. 9¾″. (Repr. Ser. 55)

Genetic parameters, their estimation and significance; by R. E. Comstock & H. F. Robinson. *Proc. 6th Int. Grassland Congr.* (1954). [8p.] 9¼″. (Repr. Ser. 56)

Relation between heterozygosis and performance in maize; by J. C. Sentz, H. F. Robinson & R. E. Comstock. *Agron. J.* (1954), **46**, 514–520. 10¾″. (Repr. Ser. 57)

Some further results in simultaneous confidence interval estimation; by S. N. Roy. *Ann. Math. Statist.* (1954), **25**, 752–761. 9¾″. (Repr. Ser. 58)

An analysis of quantitative variability in *Nicotiana tabacum*; by H. F. Robinson, T. J. Mann & R. E. Comstock. *Heredity* (1954), **8**, 365–376. 9½″. (Repr. Ser. 59)

Optimum size of sample for hand separation of forage crop mixtures into their component species in small plot experiments; by R. G. Petersen & D. S. Chamblee. *Agron. J.* (1955), **47**, 20–23. 10¾″. (Repr. Ser. 60)

An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present; by C. Clark Cockerham. *Genetics* (1954), **39**, 859–882. 9½″. (Repr. Ser. 61)

Some problems of optimum sampling; by Paul N. Somerville. *Biometrika* (1954), **41**, 420–429. 10½″. (Repr. Ser. 62)

Genetic variances in open pollinated varieties of corn; by H. F. Robinson, R. E. Comstock & P. H. Harvey... *Genetics* (1955), **40**, 45–60. 9½″. (Repr. Ser. 63)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Further contributions to the theory of paired comparisons; by M. G. Kendall. *Biometrics* (1955), **11**, 43–62. 9″. (Repr. Ser. 64)

Estimation of the parameters of skewed distribution by linear systematic statistics; by A. E. Sarhan. *J. Amer. Statist. Ass.* (1955), **50**, 196–208. 9″. (Repr. Ser. 65)

The effects of various levels of sodium citrate glycerol, and equilibration time on survival of bovine spermatozoa after storage—79°C.; by R. C. Cragle, R. M. Myers, R. K. Waugh, J. S. Hunter & R. L. Anderson. *J. Dairy Sci.* (1955), **38**, 508–514. 7p. 9¾″. (Repr. Ser. 66)

The efficiency of tests; by Wassily Hoeffding & Joan Rosenblatt. *Ann. Math. Statist.* (1955), **26**, 52–63. 9¾″. (Repr. Ser. 67)

On the distribution of the sample median: by John T. Chu. *Ann. Math. Statist.* (1955), **26**, 112–116. 9¾″. (Repr. Ser. 68)

Rules of thumb for determining expectations of mean squares in analysis of variance; by E. F. Schultz. *Biometrics* (1955), **11**, 123–135. 9″. (Repr. Ser. 69)

Analysis of confounded factorial experiments in single replications; by F. E. Binet & R. T. Leslie ... S. Weiner ... R. L. Anderson. *N.C. Agric. Exp. Sta. Techn. Bull.* 113, 1955. 64p. 9″. (Repr. Ser. 70)

The extrema of the expected value of a function of independent random variables; by Wassily Hoeffding. *Ann. Math. Statist.* (1955), **26**, 268–275. 9¾″. (Repr. Ser. 71)

Estimates of genetic and environmental variability in soybeans; by Herbert W. Johnson, H. F. Robinson & R. E. Comstock. *Agron. J.* (1955), **47**, 314–318. 10¾″. (Repr. Ser. 72)

Estimation of the mean and standard deviation by order statistics. Part II; by A. E. Sarhan. *Ann. Math. Statist.* (1955), **26**, 505–511. 9½″. (Repr. Ser. 73)

O'Brien, Terence H. Civil defence. London, H.M.S.O., & Longmans Green, 1955. xvii, 729p. maps (1 fold.) 9½″. 37*s.* 6*d.* (History of the Second World War, U.K. Civil Series)

Page, E. S. A test for a change in a parameter occurring at an unknown point. *Biometrika* (1955), **42**, 523–527. 10¾″.

Petersen, William. Planned migration: the social determinants of the Dutch-Canadian movement. Berkeley, University of California Press, 1955. London, Cambridge University Press. x, 273p. 9¼″. $3.50.

Political and Economic Planning. World population and resources: a report by P.E.P. London, P.E.P. (distributed by Allen & Unwin), 1955. 339p. 8½″. 30*s.*

Pollak, L. W. Eight-place supplement to harmonic analysis and synthesis schedules for three to one hundred equidistant values of empiric functions. 2nd ed. Dublin, Dublin Institute for Advanced Studies, 1954. 3v. 13¼″. 20*s.* (Geophysical Memoirs 1)

Rau, A. Ananthapadmanabha. Crop surveys in Mysore State 1953–54. Bangalore, Government Press, 1955. 29p. 9¾″.

Reed, Lowell J. The role of the vital statistician in the community as evolved by W. Thurber Fales. *Amer. J. Publ. Hlth.* (1955), **45**, 11–20. 9¾″.

Russell, Marion H. Diverging sex-morbidity trends in cancer of the mouth: a hospital morbidity study. *Brit. Med. J.* (1955), **2**, p. 823. 12p. map. 8½″.

Sainsbury, Peter. Suicide in London: an ecological study. London, Chapman & Hall, for the Institute of Psychiatry, 1955. 116p. 9¾″. 15*s.* (Maudsley Monographs 1)

Sampford, M. R. The estimation of response-time distribution. III. Truncation and survival. *Biometrics* (1954), **10**, 531–561. 9¼″.

The truncated negative binomial distribution. *Biometrika* (1955), **42**, 58–69. 10¾″.

Schelling, Thomas C. International cost-sharing arrangements. Princeton, University, International Finance Section, 1955. [ii], 25p. 9″. (Essays in International Finance 24)

Sovani, H. V. & Dandekar, Kumudini. Fertility survey of Nasik, Kolaba and Satara (North) districts. [Poona], Gokhale Institute of Politics and Economics, 1955. xvi, 167p. maps. 9¾″. 9*s.* (Publication 31)

Staples, Robert, *editor.* "Taxation" manual; compiled by barristers and experts under the direction of Robert Staples. 8th ed. London, Taxation Publishing Co., 1955. xx, 441p. 8½″. 25*s.*

Stephenson, William. The study of behavior: Q-technique and its methodology. Chicago, University Press, 1953. London, Cambridge University Press. x, 376p. 9″.

Syski, Ryszard. Analogies between the congestion and communication theories. London, Hivac Ltd., 1955. ii, 50p. 13″. Prepared for the Congress on the Application of the Theory of Probability in Telephone Engineering, Copenhagen, 1955.

Taylor, Wallis & Dauncey, Marjorie. Changing pattern of mortality in England and Wales: III. Mortality in old age. (a) Picture disclosed by annual specific rates. *Brit. J. Prev. Soc. Med.* (1955), **9**, 162–168. 9¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Tippett, L. H. C. Statistics. 2nd ed. London, Oxford University Press, 1956. vi, 224p. 6¾″. 7*s.* 6*d.* (Home University Library)

Wiener, Norbert. Time and organization. University of Southampton, 1955. 16p. 8½″. (Fawley Foundation Lecture)

Wilcoxon, Frank. Some rapid approximate statistical procedures. Revised edition. New York, American Cyanamid Company, 1949. 16p. fold. diagr. 8½″.

Williams, R. E. The electronic office. London, Gee & Co., 1955. 63p. 8 plates. 8½″. 15*s.*

Wise, M. E. The effect of rounding off in samples. *Statistica* (1954), **8**, 169–178. 10″.

Formulae relating to single-sample inspection by attributes. *Philips Res. Rep.* (1955), **10**, 97–112. 16p. 9¼″.

Wishart, John. Field trials II: the analysis of covariance. Cambridge, Commonwealth Bureau of Plant Breeding and Genetics, 1950. 35p. 9¾″. 3*s.* 6*d.* (Tech. Communication 15)

Wishart, John & Sanders, H. G. Principles and practice of field experimentation. 2nd ed. Cambridge, Commonwealth Bureau of Plant Breeding & Genetics, 1955. vii, 133p. 8½″. 21*s.* (Tech. Communication 18)

Woytinsky, W. S. & Woytinsky, E. S. World commerce and governments: trends and outlook. New York, The Twentieth Century Fund, 1955. lii, 907p. illus. 10″. $10.

Wrighton, Raymond. Statistical experimentation in therapeutics. *Brit. J. Prev. Soc. Med.* (1955), **9**, 135–139. 9¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

BANK OF ENGLAND

Pursuant to the Act 7th and 8th Victoria, cap. 32 (1844) *18th and 19th Geo.* 5, *cap.* 13,

(000's omitted)

1	2	3	4	5	6	7	8
		Issue Department				Collateral Columns	
Liabilities	Dates (Wednesdays)	Assets					
Notes Issued		Govt. Debt (£11,015) and Govt. Securities	Other Securities	Gold Coin and Bullion	Coin other than Gold	Notes in Circulation	Bank Rate
£		£	£	£	£	£	
1,775,362	Jan. 5	1,771,236	752	362	3,012	1,731,548	3%
1,775,362	„ 12	1,771,183	806	362	3,011	1,710,869	
1,725,362	„ 19	1,721,179	806	362	3,015	1,677,755	
1,725,362	„ 26	1,721,181	806	362	3,013	1,664,947	3½%
1,725,362	Feb 2	1,721,181	806	362	3,013	1,659,759	
1,725,362	„ 9	1,721,234	756	362	3,010	1,659,255	
1,725,362	„ 16	1,721,236	749	362	3,015	1,656,393	
1,725,362	„ 23	1,721,173	813	362	3,014	1,658,946	
1,725,362	Mar. 2	1,721,244	745	362	3,011	1,668,912	4½%
1,725,361	„ 9	1,721,235	755	361	3,010	1,678,040	
1,725,361	„ 16	1,721,247	741	361	3,012	1,680,275	
1,725,361	„ 23	1,721,268	720	361	3,013	1,683,281	
1,725,361	„ 30	1,721,232	754	361	3,014	1,696,260	
1,750,361	Apr. 6	1,746,265	721	361	3,015	1,723,750	
1,750,361	„ 13	1,746,219	767	361	3,014	1,729,958	
1,750,361	„ 20	1,746,267	720	361	3,013	1,723,954	
1,750,360	„ 27	1,746,204	784	360	3,012	1,731,522	
1,775,360	May 4	1,771,182	803	360	3,015	1,742,717	
1,775,360	„ 11	1,771,205	780	360	3,014	1,740,238	
1,775,361	„ 18	1,771,222	767	361	3,011	1,735,607	
1,775,361	„ 25	1,771,283	705	361	3,012	1,752,216	
1,775,361	June 1	1,771,182	806	361	3,012	1,765,295	
1,825,361	„ 8	1,821,172	813	361	3,015	1,778,676	
1,825,361	„ 15	1,821,217	770	361	3,013	1,787,514	
1,825,362	„ 22	1,821,210	778	362	3,012	1,785,603	
1,825,362	„ 29	1,821,280	707	362	3,013	1,782,692	
1,825,362	July 6	1,821,339	649	362	3,012	1,795,704	
1,825,362	„ 13	1,821,282	706	362	3,011	1,815,399	
1,875,362	„ 20	1,871,183	803	362	3,014	1,845,016	
1,875,361	„ 27	1,871,196	791	361	3,013	1,860,865	
1,875,361	Aug. 3	1,871,271	718	361	3,011	1,857,844	
1,875,361	„ 10	1,871,240	750	361	3,010	1,834,603	
1,825,362	„ 17	1,821,282	706	362	3,012	1,808,940	
1,825,362	„ 24	1,821,237	753	362	3,011	1,788,912	
1,825,362	„ 31	1,821,312	673	362	3,014	1,782,250	
1,800,362	Sept. 7	1,796,281	707	362	3,011	1,781,013	
1,800,361	„ 14	1,796,326	659	361	3,014	1,775,713	
1,800,361	„ 21	1,796,284	702	361	3,014	1,767,358	
1,800,361	„ 28	1,796,209	779	361	3,012	1,766,441	
1,800,361	Oct. 5	1,796,267	720	361	3,012	1,769,503	
1,800,361	„ 12	1,796,201	784	361	3,015	1,767,987	
1,800,361	„ 19	1,796,299	688	361	3,013	1,764,296	
1,800,360	„ 26	1,796,258	732	360	3,010	1,765,463	
1,800,359	Nov. 2	1,796,199	789	359	3,012	1,770,270	
1,800,359	„ 9	1,796,282	706	359	3,013	1,772,216	
1,800,359	„ 16	1,796,202	787	359	3,012	1,777,718	
1,800,359	„ 23	1,796,276	713	359	3,010	1,780,534	
1,850,359	„ 30	1,846,216	772	359	3,012	1,795,659	
1,850,359	Dec. 7	1,846,278	708	359	3,013	1,828,380	
1,900,359	„ 14	1,896,255	734	359	3,011	1,870,592	
1,900,359	„ 21	1,896,308	680	359	3,012	1,880,804	
1,900,359	„ 28	1,896,266	719	359	3,015	1,889,622	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

WEEKLY RETURN

and 9th and 10th Geo. 6, cap. 74, for Wednesday in each Week, during the year 1955

(000's omitted)

9	10	11	12	13	14	15	16	17	18
BANKING DEPARTMENT									
Liabilities				DATES (Wednesdays)	Assets				Totals of Liabilities and Assets
Capital (£14,553) and Rest	Public Deposits†	Bankers' Deposits	Other Deposits		Govt. Securities	Discounts and Advances	Other Securities	Reserve (Notes and Coin)	
£	£	£	£		£	£	£	£	£
18,203	29,792	281,514	70,628	Jan. 5	273,282	65,848	14,840	46,167	400,137
18,224	23,731	262,496	65,376	„ 12	272,412	15,708	14,863	66,843	369,826
18,308	14,397	268,107	66,075	„ 19	281,892	20,148	14,900	49,947	366,887
18,309	21,032	251,900	85,916	„ 26	283,002	16,378	15,016	62,761	377,157
18,316	18,797	261,219	65,841	Feb. 2	264,642	16,493	15,082	67,956	364,173
18,349	19,828	267,465	65,881	„ 9	272,302	15,988	14,747	68,487	371,524
18,392	18,435	282,906	63,355	„ 16	288,747	8,113	14,885	71,343	383,088
18,428	16,039	271,129	63,672	„ 23	233,272	52,350	14,857	68,789	369,268
18,481	19,329	266,272	64,141	Mar. 2	269,107	25,085	15,199	58,833	368,224
18,495	22,092	276,938	63,535	„ 9	288,947	26,945	15,461	49,708	381,060
18,509	21,210	274,574	65,129	„ 16	297,802	18,524	15,629	47,468	379,422
18,528	22,191	271,332	65,375	„ 23	308,757	8,585	15,583	44,501	377,426
18,549	15,793	275,712	66,429	„ 30	312,682	16,120	16,162	31,520	376,483
17,712	15,893	272,184	65,888	Apr. 6	308,442	20,115	14,098	29,022	371,676
17,713	20,499	274,330	64,083	„ 13	331,297	8,785	13,723	22,820	376,625
17,766	22,638	276,738	64,722	„ 20	303,462	35,770	13,805	28,828	381,864
17,822	23,201	253,889	63,621	„ 27	305,357	18,035	13,883	21,258	358,533
17,858	26,749	244,614	63,126	May 4	283,462	20,090	13,738	35,057	352,346
17,871	25,018	251,458	63,221	„ 11	276,897	29,166	13,962	37,543	357,568
17,934	19,485	257,240	65,396	„ 18	287,922	16,126	13,822	42,186	360,056
17,961	20,522	240,464	65,627	„ 25	293,217	11,801	13,988	25,568	344,574
17,995	22,248	239,520	66,434	June 1	308,967	10,801	13,948	12,482	346,197
18,002	16,549	248,407	79,464	„ 8	279,242	20,081	14,007	49,092	362,421
18,047	21,638	251,042	70,236	„ 15	289,677	16,951	14,079	40,255	360,962
18,103	26,304	245,033	66,317	„ 22	290,377	7,176	16,036	42,169	355,757
18,150	14,706	246,602	72,910	„ 29	285,457	5,725	16,117	45,069	352,368
18,171	19,194	247,066	68,026	July 6	274,922	29,556	15,926	32,052	352,456
18,259	13,118	255,127	67,899	„ 13	298,207	27,656	16,184	12,357	354,404
18,265	23,777	245,443	67,496	„ 20	278,257	28,317	15,669	32,741	354,983
18,307	15,782	240,831	68,838	„ 27	290,202	20,957	15,709	16,891	343.758
18,334	27,797	216,450	72,353	Aug. 3	289,562	9,832	15,646	19,895	334,934
18,369	26,295	230,060	68,100	„ 10	257,557	26,457	15,679	43,132	342,824
18,406	21,524	224,720	68,677	„ 17	289,017	9,547	15,961	18,803	333,327
18,434	14,451	233,011	68,435	„ 24	249,232	30,207	16,056	38,836	334,330
18,463	13,639	243,593	69,994	„ 31	263,357	20,017	16,840	45,476	345,689
18,469	13,110	253,292	65,814	Sept. 7	286,534	25,456	16,958	21,738	350,686
18,495	15,021	257,241	66,055	„ 14	277,505	35,066	17,168	27,072	356,811
18,523	15,904	263,066	65,956	„ 21	278,405	32,186	17,435	35,424	363,449
18,551	17,070	239,321	68,232	„ 28	278,575	11,055	17,190	36,353	343,173
17,705	21,700	250,710	66,972	Oct. 5	270,575	36,730	16,491	33,291	357,087
17,743	17,970	257,937	65,895	„ 12	280,820	27,340	16,584	34,801	359,545
17,778	13,120	259,211	66,307	„ 19	271,115	30,190	16,619	38,491	356,416
17,838	24,482	250,630	65,685	„ 26	284,240	20,150	16,919	37,325	358,634
17,847	17,169	251,783	66,261	Nov. 2	268,515	35,340	16,694	32,511	353,060
17,869	18,743	246,640	65,606	„ 9	261,800	39,685	16,813	30,560	348,858
17,910	16,242	252,040	65,287	„ 16	295,170	14,480	16,770	25,059	351,479
17,990	13,085	255,840	66,357	„ 23	294,875	19,210	16,951	22,236	353,272
17,994	13,086	263,059	71,168	„ 30	270,665	20,820	16,738	57,084	365,307
18,023	14,484	249,299	69,586	Dec. 7	287,915	22,155	16,991	24,331	351,392
18,039	15,194	239,846	69,874	„ 14	281,530	12,155	17,164	32,105	342,954
18,147	21,991	221,321	71.040	„ 21	282,355	11,085	17,176	21,883	332,499
18,149	15,275	245,183	71,740	„ 28	282,690	37,720	16,884	13,053	350,347

† Including Exchequer, Savings Banks, Commissioners of Nat. Debt., and Div. Accounts.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Vol. 119.] [Part III

Journal of the Royal Statistical Society

SERIES A (GENERAL)

PART III, 1956.

United Kingdom Indices of Wholesale Prices, 1949–1955

By H. S. Phillips, B.A., Ph.D.

[Read before the Royal Statistical Society, March 21st, 1956,
Professor A. Bradford Hill, C.B.E., F.R.S., Honorary Vice-President, in the Chair]

Summary of Contents

Paragraphs.

Introduction 1–3

Part I: A Description of the New Scheme of Board of Trade Wholesale Price Index Numbers (June 30th 1949 = 100) and of the Practical Problems Surmounted in Constructing and Maintaining them.

(A) The new scheme briefly described (with Table A) 4–5
(B) The preliminary approach to the task of collecting prices 6
(C) The collection and processing of prices and the subsequent calculation of index numbers 7–8
(D) Problems involved in determining the weights to be used in constructing price index numbers 9–19
(E) The problem of discontinuous series 20–27
(F) The problem of quality changes 28–34
(G) The presentation and form of publication of price indices 35–37
(H) The ending of the old (1930 = 100) wholesale price series 38–39
(I) Rebasing of the present wholesale price series on 1954 = 100 40–41

Part II: The Relative Movements of the Main Wholesale Price Index Numbers in the New (June 30th, 1949 = 100) Series: Some Analyses and Comparisons.

(1) A selection of the published index numbers, 1949–1955: with Tables 1a and 1b . 42–43
(2) Simplifying the construction of major composite index numbers; with Table 2 . 44
(3) A comparison of the price movements of related groups of increasingly processed commodities; with Table 3 45
(4) The relative volatility of the prices of selected commodities; with Table 4 . . 46
(5) Comparisons between wholesale and retail price indices; with Charts 1 and 2 and Table 5 47–48
(6) A comparison of various published price indices for raw materials; with Table 6 . 49
(7) The reliability of the Laspeyres formula; with Table 7 50

Appendix I: The questions asked and the problems met in the process of obtaining the co-operation of industry in the collection of prices . . . . . (1)–(17)

Appendix II: The storing and processing of price information and the calculation of price relatives and index numbers . . . . . (i)–(v)

Introduction

1. It was realized before the Second World War that the old (1930 = 100) series of Wholesale Price Index Numbers was in need of revision. Although these indices could not be reconstructed during the war years that followed, thought was given to the need for something much more useful

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and meaningful than a revised series of index numbers based on the old idea. The need for an entirely new and more comprehensive system of wholesale price indices, based on a more useful concept and providing price guidance on each of the main sectors of economic activity, covering manufactured products as well as materials, was set out by Mr. Stafford in his paper to this Society on May 23rd, 1951, and the new (June 30th, 1949 = 100) series of Board of Trade Wholesale Price Indices was introduced in the *Board of Trade Journal* almost simultaneously.

2. It has been possible, in the period of nearly five years since that paper set out the basic concepts on which the new scheme would be based, to come close to the completion of the new complex of indices. While some of these indices, including important ones in the Engineering sector, have not been published, nearly all the indices required to complete the scheme are now calculated and in use. This, therefore, seems to be an appropriate time to produce a further paper which not merely surveys the practical and technical problems met and faced in completing the new series of wholesale price indices, but also provides some analysis of their various movements. That is what I attempt to do in Parts I and II of the paper that follows.

3. I wish to acknowledge the assistance given by Miss Helen Morgan in the preparation of this paper and to emphasise that no such paper would have been possible but for the team work of all those who, under the direction of Mr. Stafford, have been engaged on wholesale price work in the Board of Trade over the past six years, or without the co-operation of statisticians in other Government Departments.

PART I

A DESCRIPTION OF THE NEW SCHEME OF WHOLESALE PRICE INDEX NUMBERS (JUNE 30TH, 1949 = 100) AND OF THE PRACTICAL PROBLEMS SURMOUNTED IN CONSTRUCTING AND MAINTAINING THEM

(A) *The New Scheme briefly described*

4. A summary of the scheme towards which we are working in the Board of Trade may be seen at Table A. Briefly we aim at preparing sector input, output and "consumption" indices, together with as many commodity indices as may be useful under each of these heads. In general, it is the primary objective of the Board of Trade to gather together a good deal of commodity price information, in the belief that individual commodity price indices, not being so dependent on weighting for their construction but capable of rapid and *ad hoc* combination in different sector patterns and proportions, for many different purposes, are the really useful ones. The usefulness of sector composite indices, on the other hand, is limited by the extent of the demand for the index for any particular sector covered and by the out-of-dateness of the base weights necessarily used in their construction so that, while providing invaluable general economic indicators, they may for many purposes be less useful than their component parts. The entire scheme, in its present virtually complete state, is based on several hundred price series for materials and more than 5,000 for manufactured products. These prices are contributed, directly, by some 1,500 firms and 120 trade associations (representing a further large number of manufacturers) and they are combined into approximately 700 distinct sector or commodity group indices, some 95 per cent. of which are base weighted using 1948 sales or purchases (as appropriate) for weights. Almost all the indices for materials and the majority of indices for manufactured products are monthly ones and reflect average daily prices in each month; the exceptions—mostly relating to items the prices of which change less frequently—are based on prices at a given quarterly date. Something like 90 per cent. of the indices for materials and 60 per cent. of those for manufactured products are either published monthly in the *Board of Trade Journal* or are available to the public on inquiry.

5. The fact that to-day, some years after our work was begun, a number of quite important indices among those referred to in Table A cannot yet be published, even though most are available to the Government, stresses the considerable practical problems (which this paper seeks to record) that have to be faced in collecting prices and calculating index numbers. It must be emphasized that this price collection work has all been done on the basis of voluntary relations with industry and that—both in the Board of Trade and in the other Government Departments involved—the staff available for collecting prices, maintaining them in good order and calculating indices has been very limited. In the face of these facts we have felt that the description of our methods, and the problems we have had to deal with, would be useful not only to those in this country who use our

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

index numbers and want to know more about them, but to statisticians, economists and officials in other countries who are thinking along our lines and might learn from our mistakes and experience.

TABLE A

Summary Programme for the Publication of Wholesale Price Indices in the United Kingdom

(1) INPUT INDICES*

(*a*) *Sector Indices for materials entering into industry*

Materials used in home agriculture.
Basic materials (including fuel) used in manufacturing industry.
Basic materials (other than fuel) used in non-food manufacturing industry.
Basic materials (other than fuel) used by food manufacturing industries.
Mechanical engineering materials, other than fuel.
Electrical engineering materials, other than fuel.
Textile industry materials, other than fuel.
Building and civil engineering materials.
House-building materials.
All fuel used in manufacturing industry.

(*b*) *Commodity Indices for materials entering into industry*

(2) OUTPUT INDICES†

(*a*) *Sector Indices for output (including exports) of industry*

Output of home agriculture.
Products of non-food non-fuel manufacturing industry.
Manufactured foodstuffs.
Iron and steel products.
Mechanical engineering products.
Electrical engineering products.
Textile products.
Clothing.
Footwear.
Chemicals and allied products.
Timber products.
Paper products.
Fuel products.

(*b*) *Indices for individual commodities and groups of commodities produced in the United Kingdom*

So far as possible all individual industry headings included in the United Kingdom Standard Industrial Classification to be represented by the commodities for which price indices are available. The commodities include semi-manufactures which do not enter into the output of the sectors for which indices are published unless exported. (Commodities of which there are significant imports and which enter into home consumption may be excluded and shown only at 3 (*b*).)

(3) CONSUMPTION INDICES.

(*a*) *Sector Indices for groups of commodities entering into final consumption (including imports and excluding exports)*

(The precise form to be taken by these indices is not yet certain)

(i) Capital goods including:
Plant and machinery.
Buildings, and certain classes of vehicles.

(ii) Consumer goods including:
Manufactured personal consumers goods other than fuel, food, drink and tobacco.
Food (including output of home agriculture, imported food and food manufactured at home).
Drink and tobacco.

(*b*) *Indices for individual commodities entering final consumption.* (Excluding those wholly manufactured at home shown at 2 (*b*))

* See also Part II, Table 1(*a*).
† See also Part II, Table 1(*b*).

(B) *The Preliminary Approach to the Task of Collecting Prices*

6. The procedure adopted before making any approach to industry for prices might be summed up as follows :

(i) The whole range of production was first divided along the lines of the Standard Industrial Classification between those Departments responsible for the activity of each

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

industry, a responsibility which in 1948/1949 had a very positive meaning—since there were controls over most commodities either directly or through the rationing of materials. In some cases the Board of Trade, as the Department publishing the indices, took over responsibility for collecting prices for commodities which were, strictly speaking, the production concern of other departments. In the outcome it could be said that the task of collecting prices for Order I commodities (agriculture, forestry and fishing) rested, naturally, with the Ministry of Agriculture and Fisheries (now the Ministry of Agriculture, Fisheries and Food); for coal, coke, oils, gas, electricity, etc. (parts of Orders II, IV and XVIII) with the Ministry of Fuel and Power ; for Orders VI, VII and parts of VIII and IX (i.e. the main metal-using industries) with that part of the Ministry of Supply recently transferred to the Board of Trade; for Order XIII (food and drink) commodities with the Ministry of Food (now the Ministry of Agriculture, Fisheries and Food) and for Order XVIII (building and contracting with the Ministry of Works. Collection of prices for the remaining items, namely those in Orders II, III, IV, V, X, XI, XII, XIV, XV and XVI fell to the Board of Trade, who also had a general co-ordinating responsibility for the preparation of composite indices and for the collection of prices and the preparation of indices for materials used throughout manufacturing industry.

(ii) The main materials and manufactured goods for which price information would be required were then determined by those with detailed knowledge in the Departments concerned, by reference to the 1935 Census of Production, current short term production returns and any preliminary information that was to hand from the 1948 full Census of Production.

(iii) The question next arose how far it would be wise to rely on published prices from journals or from market reports. So far as most "materials" with a recognized market price were concerned, there was little doubt that published prices would prove reasonably reliable sources of information. Any error involved in using them was considered much too small to be regarded as significant, in relation to the unpopular and onerous procedure of obtaining prices from several hundred individual trade sources, each recognizing the market price and tending to refer one to it. It is clear that in collecting prices and measuring the price movements of raw materials "perfectionism" would make almost any price collection system unworkable and structurally correct indices would, in the outcome, perhaps differ only slightly from those produced by using published market prices. It is well known that in most of the raw material markets there tends to develop the quotation of a "basis" price and that around this price there are, in fact, a large number of real differences, flowing from quality, size, shape and other kinds of premium, the relationships of which may change one to another slightly if not significantly, from time to time, but which, for better or worse, have normally got to be ignored in calculating price indices. There are, of course, important qualifications one must make to this idea of the general acceptability of recognized market prices. There are occasions, for example, when prices for materials regularly appear in journals as being "delivered" to particular areas. In these cases it is dangerous to use published prices without carefully considering possible alternative sources of information, since the *movement* of prices "delivered (say) London", particularly for building materials, can be widely different from the movement of prices, on average, "delivered (say) nearest railway station anywhere in the United Kingdom". One must, of course, also watch carefully to see whether (as in the case of copper from May 9th, 1955) the recognized market is temporarily by-passed and ceases to reflect average prices adequately. This kind of exception has detracted from the once widely accepted price leadership authority of some markets, and with less hedging and more fixed price (and strategic stock) buying, the reality of prices quoted on markets is perhaps less certain than it used to be for spot and cash business. These considerations are making the construction of reliable price indices for a few key materials, often dealt with only in marginal quantities on recognized markets, more hazardous. These difficulties can be discussed, as they arise, with representatives of the trades concerned and if the recognized market for a particular commodity appears no longer to represent the real price of that commodity then additional sources of information can be sought.

(iv) During the period of controls many prices were extracted from Statutory Instruments. Latterly, the number of controlled prices has become limited to iron and steel scrap and a few food and fuel items but for most iron and steel semi-manufactures, the Statutory

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Instruments have, in effect, been replaced by the Price Determination Lists issued by the Iron and Steel Board.

(v) With very few exceptions prices for manufactured products cannot be obtained from journals, or if they can be so obtained are not sufficiently representative for purposes of calculating index numbers. For prices of manufactured products it was thus necessary for us to decide what trade associations should be approached and to go to these bodies, explaining what we were doing, seeking their advice and their co-operation, and asking for the names of the firms from whom representative prices might be collected. The importance of approaching trade associations lies not in the expectation that they have all the answers about prices, but rather that they are generally in a position to recommend suitable firms to approach for prices. Even trade associations who do not concern themselves with prices are (with a few exceptions) quite prepared to give advice on some sources of price information within the trade and there can be little doubt that a preliminary approach to associations greatly increases the goodwill subsequently shown by member firms within each trade. In many cases trade associations also provide technical information invaluable in the subsequent approach to firms. Our experience has been that the approach to trade associations thus produces a co-ordinated rather than a haphazard and possibly over-lapping approach to individual manufacturers.

(C) *The Collection and Processing of Prices and the subsequent Calculation of Index Numbers*

7. The successful collection of a sample of prices sufficiently adequate to enable reliable price indices to be constructed depends on the voluntary co-operation of industry, and the practical problems involved in obtaining this assistance on a wide scale are numerous. I do not wish to burden the main body of this paper with details of all the innumerable interesting and important questions that arise in the process of approaching manufacturers for prices, in outlining the objects of index numbers to trade associations or in explaining how the indices will be constructed. I should, however, record that one cannot succeed in building up a complex system of price indices without answering these practical questions clearly and continuously. In the belief that the experience gained by the Board of Trade in this field will be valuable to those connected with this kind of work, I have embodied at *Appendix I* a fairly complete survey of all the various questions that have to be asked and answered and of the kind of approach that has been adopted in consulting industry.

8. Having obtained the basic price data required for the construction of the indices it is essential to store and process this information carefully and methodically, devising self-correcting procedures for calculating index numbers and maintaining their reliability. Some of the various mechanical operations and procedures involved in coping with the mass of wholesale price data maintained by the Board of Trade are described at *Appendix II*. It is hoped that this description will be of special value to those who may be facing this kind of problem for the first time.

(D) *Problems Involved in Determining the Weights to be Used in Constructing Price Index Numbers*

9. One of the major tasks that has to be faced in constructing any wide system of index numbers is the calculation of suitable weights. In carrying out this task it is necessary first to assess the weighting significance to be given to each of the commodity price series that are available and, secondly, to decide how to allocate the weights appropriate to the various price series that are not available.

10. Although there are no serious conceptual difficulties in the first task, which mainly involves extracting input and output value figures from Census of Production reports, the magnitude of this part of the work can be gauged by appreciating that, in a scheme like that on which the Board of Trade embarked in 1949, one is dealing with several hundred different materials and several thousand products. An attempt has been made in the Board of Trade scheme to collect price information for every specifiable item which, on the basis of Census data, would carry a weight of 1 per cent. or more (or, if less, with a sales or purchase value of over £1 million) in any input or output Standard Industrial Classification sector down to a minimum list heading. In this scheme it has also been our constant endeavour to see that the sum of the values of specifiable items neglected in any sector is always less than 5 per cent. of the total value of specifiable items included. When estimating the value

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of particular "Materials Purchased" (and products sold) by different industries it has to be remembered that the "Materials Purchased" (and "Sales of Principal Products") tables in the Census of Production reports often contain substantial "basket" totals of items "not elsewhere specified"; lack of adequate knowledge about the breakdown of such totals may, in some cases, make the resultant estimated weights liable to a significant margin of error. Where the Census of Production does not provide a sufficient breakdown of information, which is often unavoidable in the case of manufactured products, it is necessary to make *ad hoc* estimates with the help of the trades concerned. There are also cases, as for example that of "wooden containers", where it is necessary, in order to obtain even approximate output weights for different classes of container, to search through the "materials purchased" sections of the Census reports on trades using wooden containers. This illustration is only one of many which could be quoted as suggesting that really detailed input/output analyses would help those concerned with determining the weights to be used in constructing price indices.

11. This kind of extraction is primarily a matter of hard work, but since our indices are what the United Nations Statistical Commission describes as "net" (as distinct from "gross") weighted sector indices, excluding as they do all transactions within sectors, care must be taken to avoid double-counting or "duplication". The extent to which particular products are transferred within the same trade (or sector of industry) has to be assessed from the figures available. Allowances must also be made for "business usage" in calculating weights for indices for those personal consumer goods which, like passenger cars, writing paper or carpets, also enter into business consumption on a large or significant scale.

12. Far more complicated than the task of determining the appropriate weights for commodities that are represented by actual price quotations in an index is the second task of deciding what to do about the weights of price series that are not available. This question divides itself into two somewhat distinct parts—that concerning "commodities" not covered by price quotations at all and that concerning commodities for which the "price quotations" available represent only a very small sample of all those that might have been obtained. Of these two allied weighting problems the first is the least difficult. It arises from the fact that in any wide system of index numbers there are bound to be a certain number of commodities either too unimportant for prices to be worth collecting or for which prices cannot be obtained. In each case one has to decide whether the value weight of the missing commodity should be imputed to some other commodity, or to some group of commodities, making the decision on the basis of the evidence as to which of these commodities or groups of commodities is best qualified to represent the price movement of the missing commodity. When the missing item is, owing to inevitable sampling, no more than a variant on one already included in the index, there being clear evidence that its price movement will be similar, there is no problem. In the absence of any such evidence one may be forced to impute the value of the missing commodity to the whole of the rest of the group index in which it should feature.

13. One of the main difficulties in imputing the weights of missing commodities to those represented is lack of technical knowledge. This is especially true in those special cases where there is a combination of price movements for several different types of product, and where the particular item missing may well be suited to imputation to one or other of these but not to all combined—a fact which may not be sufficiently appreciated by the person making the judgement. For an example of the possible difficulties of imputation one may look at the index for Domestic Electrical Appliances and note that the mid-1955 index numbers for cookers, heating appliances and vacuum cleaners range from 116 to 123 (June 30th, 1949 = 100), whereas the index numbers for washing machines and refrigerators, the demand for which has been rapidly increasing, were at 103 and 100 respectively. Features such as new inventions, the extent of export demand, the strength of price fixing agreements within the trade, recent developments in the production of other commodities, etc., all affect the price pattern quite considerably. Because of this there are a number of decisions on imputation which cannot wisely be taken without consultation with those closely connected with the technical side of the production and with the marketing of the commodities concerned; in the case of an index like that for Domestic Electrical Appliances it is clear that if data on prices of some group of items were not available—then any imputational decision taken without technical advice might lead to serious error. Such advice is also essential if, for example, one is deciding how to impute the weight for certain "tailor made" products (for which a regular series of comparable prices cannot be obtained) to the prices of very similar but "standard" products.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

14. It is perhaps useful to give a specific example which lends support to the view that incorrect decisions on imputation, so far as missing commodities with small weights are concerned, do not usually prove significant. Such an example can be provided by looking at the price index for "Clothing and Footwear". Rubber footwear items carry a weight of 2 per cent. within this index and in the (assumed) absence of price information for these items it might be decided to impute their weight to the index as a whole. With equally good reason it might perhaps be decided to impute this weight to that for leather footwear. Perhaps again it might be decided to give the weight to a "dummy" index brought in from outside, for example that for "rubber goods other than tyres and tubes". Each of these three decisions would produce a different answer, since the level of the clothing and footwear index at mid-1955 was 113 compared with 122 for footwear (other than rubber), 131 for rubber footwear itself, and 146 for "rubber goods other than tyres, tubes and footwear". These differences were even more marked during the period of peak prices at the end of 1951: the latter two indices, for example, then standing at 139 and 171 respectively. Since the weight involved is only 2 per cent. it is clear that whatever decision was taken the index would never have been affected by as much as one whole point—and if the index for "other rubber goods" had been in use at mid-1955 by only 0·3 points, despite the substantial differences shown above. The example serves, nevertheless, to illustrate the danger of assuming that the best decision involves imputation to items made from the same material, when, in fact, other economic factors involved may be as, or more, important (in this case the competitiveness of certain classes of leather and rubber footwear and the similar technical conditions of their manufacture, as, for example, in the matter of labour costs).

15. It should be made clear that, although the above kind of imputational decision tends, in practice, to have no significant effect, this is only because the unrepresented items can usually be kept down to a minimum and thus carry only a small weight. Should a major omission be made in collecting prices for any index, this could involve imputational decisions which would have a decidedly significant effect on the index. Certainly the problems of imputation are sufficiently real for one to wish to avoid them as much as possible. Thus, if one considers the index for " toys and games", in which goods made from a variety of materials are represented, one finds that "wooden" items carry a weight of 8·5 per cent. Supposing that it had proved impossible to obtain price quotations for these items (which was not the case) it might have been necessary to impute their weight to the index as a whole, since wooden toys could not be thought to have anything in common with any one of the sub-indices, which include metal toys, soft toys and dolls, plastic toys, etc. In fact, at mid-1955 the wooden toys sub-index—which stood at 144—was 44 points higher than the plastic toys sub-index (100), and 23 points higher than the whole index (121) which includes it. In this case, therefore, a decision to impute the wooden toys' weight to the index as a whole would have tended to reduce the index by 2 full points at mid-1955.

16. Fortunately, there is much detailed technical knowledge available in industry and in the Government service which is available to the statistician when he is confronted with these problems.

17. The fact that prices for any given commodity can, in general, be collected from only a small sample of firms necessitates making weighting assumptions about the price movements of items manufactured by firms not represented in any particular index. This gives rise to the second type of imputational problem referred to, one which does not arise if prices can be obtained from all or most of the firms engaged in producing a particular item—an ideal which can, fortunately, be achieved in some trades. Where a fairly complete sample of prices can be obtained it is normally possible to allocate an appropriate weight to the price quotations received from each firm—based on knowledge of their production—and hence to avoid any question of imputation, but since the range of producers that can be covered is severely limited, both by official resources and by the response that can reasonably be expected from them, there will be a large number of commodities for which the price information available represents only a minute sample of all the prices being obtained in the sales universe with which one is concerned. While this immediately raises the theoretical question, how to allow for the missing price quotations, it should be stressed that no real problem is created unless there is evidence from the sample of prices already available that there is likely to be some appreciable dispersion in the price movements of like products made by different firms. As this dispersion is found for many products to be very slight, the extent of this sampling

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

problem can obviously be exaggerated. Where there is clearly no significant dispersion there is little point in taking more than an unweighted average of the price movements. The extreme example of this is where there is no dispersion at all between price movements, as in the case of commodities sold at one national price.

18. A real weighting problem due to small samples of price quotations for particular commodities does nevertheless exist, especially in trades where independent price policies are followed and there is a dispersion in the price movement of like items. In most of these cases there is no effective way of obtaining adequate data other than that of striving continually to improve one's sample.

19. There is little doubt that the composition of each commodity index has to be looked at on its own merits to see whether some internal weighting between the quotations received for like items from different firms is possible—on the basis of confidential production data received—or even desirable. It will, in general, be found that, either because the dispersion of price relatives is small, or because the samples themselves are too small for weighting assumptions affecting the unrepresented items to be reliable, a system of equal weighting of quotations for similar items is as good a solution as any. There will, however, be numerous and important exceptions including (*a*) the type of price index where the contributors include all or nearly all the major firms in the trade, whose relative significance one to the other can easily be reflected in some form of formal or self-weighting; (*b*) indices for commodity groups, such as shoes or furniture, where a precise distinction between quality and mass producers is essential; (*c*) indices for like items where the dispersion of price relatives is marked, and improved samples are needed.

(E) *The Problem of Discontinuous Series*

20. The problem of discontinuous series is primarily that of what action to take when the particular item for which one is receiving prices (or the firm supplying the information) goes out of production, but it arises whenever prices cease, for whatever reason—temporarily or permanently —to be available for any item featuring in an index. The treatment of some kinds of seasonal series is part of this problem.

21. It may sometimes happen, when an item is withdrawn from production, that the remaining items are considered sufficiently numerous and representative, and the dispersion between their price relatives is sufficiently small, for a replacement to be unnecessary and thus for the discontinuity to be, in effect, ignored. For example, a quotation for flannelette, one of a relatively large sample in a sub-index for flannelettes, might lapse when the contributing firm ceased production. The effect of withdrawing this item from the sub-index (i.e. imputing the weight to the sub-index and averaging the price relatives of the remaining items) might be to decrease the level of the sub-index by approximately 0·1 and thus to lower the index for all "cotton cloth" by only 0·006 on the basis of the weights currently used in the index of that title. In such a case a replacement could be considered unnecessary, imputation of the missing item to the sub-index as a whole could be permanent and discontinuity would create no problem.

22. There will also be no problem of discontinuity if the contributor of the price for the item which ceases to be available can *immediately* provide a straightforwar dreplacement himself, but if a new contributor has to be sought or if careful consideration has to be given to the statistical treatment of the link between the old and the new items thus causing delay, then there is discontinuity. It is this discontinuity, due to the interval without adequate price information, that creates the problem we are here concerned with, not the question of how to splice one price on to another, which is dealt with separately as a "quality change" problem in the next section of this paper.

23. It may sometimes take several weeks or even months to obtain suitable replacements for discontinued items, especially if correspondence with potential new contributors is involved, and in the meantime one has to decide how the price series should be maintained. If the prices of the other items in the index remain steady, the missing series can safely be maintained at its last relative until a replacement arrives. If, however, prices are fluctuating considerably, or moving very decisively in one direction, it is advisable to apply a "dummy" price movement—usually based on the average movement of the items remaining in the sub-index—and to use this dummy price movement by applying it to the last price in the missing series. The correctness of one's decision can be judged—and if necessary revised—if the manufacturer of the new item is able to provide price information back over the period in question.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

24. If a replacement proves unobtainable (the item may have gone out of fashion) one has to decide whether to give the weight of that item to a similar item which has had a fairly comparable price movement, or to drop the series from the index altogether, thus, in effect, redistributing its weight over the items remaining. This sort of problem arises where an imported item ceases to be imported altogether, as may happen with particular types of timbers, silk, etc. There are, however, cases where redistribution of weights consequent upon the removal of an item from an index would seriously distort it—because the relative of the discontinued series was exceptionally high or low—and in these cases it sometimes proves necessary to maintain a dummy series on a more or less permanent basis, until the series is revised.

25. It may be asked how one can decide without being unreasonably arbitrary, what a dummy movement for a missing item should be. In most normal cases the difficulty of the task can easily be exaggerated. One always has the experience of previous months to see whether the price of the item concerned has moved with the sub-index, in which case the solution of adopting the movement of the sub-index itself will normally be applicable, or one has some proof that the item moves in concert with the price movements of other similar items, though not necessarily with the movement of the index as a whole—and can thus inject a dummy movement similar to the related items.

26. A particular aspect of the problem of discontinuous series is *seasonality*, where a particular commodity ceases to be available or priceable over a given part of the year. Although the problem has not been very marked in the indices so far published it is considerable when dealing, for example, with fruit and vegetables and no final decisions have yet been taken about the treatment of the prices of agricultural products. It seems that most of the conceptually possible solutions are equally and inevitably artificial. One possible solution is to use variable weights from month to month, reflecting month by month consumption in the base period. Apart from the very serious difficulties in applying such a method, however, the paucity of data available for determining the monthly weights and the "hybrid" nature of the resultant index, often make it impracticable. Such a method would produce an index which was useful only as a measure of price changes in the same month from year to year and would not be meaningful from month to month of the same year. Since the basket of items being priced would not be the same from month to month, as in all the other indices in the series, the monthly movements of such an index would tend to be rather erratic. This is an important subject in itself, and if the Government does devise a satisfactory and practical scheme for dealing with the prices of seasonal agricultural commodities it may be desirable to apply that method to the indices for such commodities as cotton and wool, prices of which are also affected by harvest conditions and sales of which are significantly heavier in some months than in others.

27. The case of non-food seasonal commodities that cease to be sold widely in a particular period of the year, for example, summer dresses in winter and men's overcoats in high summer, is not really comparable with the much more difficult problem where the item not only ceases to be demanded but ceases to be available for sale and may (owing to harvest conditions) return to sale in the following season at a very different price from that at which it sold in the previous one. The movements of prices of women's dresses and winter overcoats do not tend to vary, at all violently, from those of the general family of commodities to which they belong. These garments are available during both winter and summer, since manufacturers (with whom we are concerned in this context) usually keep some "stocks", and to this extent one might dismiss the problem on the grounds that some prices are available and that these will do. This is, however, rather too placid an answer since the price quoted is probably a "special" one for very small "out-of-season" quantities. While it may be "convenient" in such cases to use the last price quoted for the commodity before it went out of season, it is difficult, especially if the remaining items in the index move substantially during this period, to defend a practice whereby seasonal items remain stable and thus, if heavily weighted, give a stability to the index which may be found to be a little artificial when the next season's price becomes available. The alternative of cutting the item out of the index altogether and transferring its weight to the remaining items, or to one of these, is also difficult to defend when it creates a sudden change in the index owing to the price relative of the seasonal item having stood at a significantly different level from that of the index as a whole or of the particular item to which its weight was applied. Accordingly, it seems that the best practical solution, at least in periods of significant price change, is to apply a "dummy" movement to the out-of-season item,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

basing this on the average movement of the remaining items in its sub-index. In general this latter solution produces reasonable results for non-food items and, partly because the out-of-season period is very short, the assumptions it involves can be revised when the new season begins. While this procedure appears to be favoured in a number of countries, I do not wish to suggest that it could be applied to perishable agricultural products. Unlike summer dresses or winter overcoats, seasonal agricultural products are liable to change in price, severally, from year to year—owing to special harvest conditions for each item, in a way which bears no necessary relationship to the changes in price of agricultural products in general. It is very difficult to devise any entirely satisfactory way of dealing with these seasonal products in a price index.

(F) *The Problem of Quality Changes*

28. The problem of allowing for quality changes in constructing index numbers is a universal and difficult one. The most common kind is that affecting commodities which continue to be available and priceable from month to month, but which may from time to time (perhaps quite frequently) be changed sufficiently in specification for this to affect their prices ; it may be that, in the case of toys, a plastic part replaces a metal one; or in a suite of furniture, that the design is changed; or, in prams, that rubber handles are removed; or, in cars and television sets, that the finish is varied; it may be in the case of matches that the number in the box is altered or in the case of beer that the gravity is changed or in manuscript books that the number of pages is increased or decreased, and so on. To ignore the change that is being made in the physical content of the item is normally to create an unnatural bias in the index which may well tend to be cumulative. The tendency for this bias to be cumulative is especially marked if raw materials are rising rapidly in price and compensatory changes are being made continuously in the technical specifications of commodities. In such cases the great advantage of obtaining prices direct from manufacturers, as is done in the Board of Trade, is that they are in the best possible position to advise as to the real cost content of the quality changes that are taking place: manufacturers are acutely conscious of all the changes they are making and (in our experience) are usually most willing to discuss them fully. In many cases of quality change adequate information is, therefore, available for a reasonably accurate assessment to be made of the change, and the price series can be adjusted upwards or downwards as the case may be.

29. Difficulties arise, however, in some circumstances. For example, the use of an additional proportion of synthetic fibres in a wool cloth or the use of slightly different dyes are things which are difficult to assess, especially in the clothing field where the manufacturer or seller of the "garment" is probably not the manufacturer of the actual material from which it is made. There is, therefore, a danger that insufficient allowance for quality changes may sometimes be made and the movement of the index may, accordingly, be damped.

30. The short answer—so far as there is any answer—to the question of how to allow for quality changes in index numbers is that allowances must normally be made by using a link provided by the price contributor. Where such a link is impossible, one has normally to transfer to the substitute commodity on the basis of an estimate of what the price of the previous item would have been had it continued on sale. Even if it has continued on sale, and there is a "price", there is a danger that small quantity "clearance" sales will be unrepresentative, but if it has not then the procedure is full of hazards, since many manufacturers will be naturally reluctant to make a hypothetical estimate of what a price might be in unknown circumstances. There is, accordingly, a danger that these links will tend to be made at the level of the previous relative. Because of this any increase or decrease that might have taken place in the real price will tend to be ignored in the index.

31. The introduction of new but not necessarily quality-reducing materials, or devices, often occurs at times when prices are going up and when innovations are being made to offset price increases. A particular aspect of such changes is that which entails changes in fashion. There are many cases where inability to assess the real cost significance of any particular quality change makes it very difficult to decide clearly whether or not it is reasonable to link one price to another at the old relative, or to treat the two prices as relating to comparable specifications and to show a straight increase or decrease, or to adopt some third procedure. So often, as in changes of pattern in furniture suites, changes in fashion of women's dresses, and changes in pattern of wireless cabinets,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

etc., there is some case for assuming that there has been no fundamental change in quality but rather a change in the tastes of buyers, and that in such cases lower (or higher) prices should be faithfully reflected in lower (or higher) index numbers. One must be very careful of this kind of argument, however, since, although it may conceivably be a realistic or fair argument, from the point of view of consumption price indices and the assessment of the cost of living, it is not reasonable in an index which is designed to measure price changes of a quantum of output. In such indices it is essential for the commodities being linked to be thought of in "physical" rather than in consumer satisfaction terms, and, where possible, guidance by the manufacturer as to the cost significance of the physical changes made is thus needed.

32. It would be fair to say that in compiling price indices of final goods the Board of Trade would not make any assumption about the "satisfaction" provided by given commodities which change in quality, but would tend to err on the side of assuming that two different things must be treated as different and that links between them should be made at the old relative if in doubt. Where a change is obviously the result of a real invention there can be little doubt that one is dealing with an improvement in quality, even should the *price* remain constant, and that an adjustment to the index to allow for this improvement does need to be made. How to make this allowance, and how much to assume the improvement to "cost", is another matter, merely illustrating that some price index numbers should really be maintained by a special class of mortals and may only be understood by them! For example, the "cost" of a radical improvement to some major product may depend (per unit) on the number of units over which any capital sum involved is estimated to be spread. Given that the amalgamation and reorganisation of two major manufacturing concerns took place simultaneously with the introduction of an improved type of product this product might be cheaper and the cost of the improvement negative.

33. The pros and cons of assuming an increase or decrease in price or of a straight splicing procedure in given situations is very fully considered by Erland v. Hofsten in his book *Price Indexes and Quality Changes* published in Stockholm in 1952, which is, however, primarily concerned with adjustments to consumer retail prices indices. I shall not try to add to what has been said in that book but would sum up my attitude by saying that every effort must be made to achieve a realistic link in every instance of quality change as it arises, each on its own merits. Human nature being what it is and staff limitations in statistical offices what they are, these various quality links will never be entirely satisfactory and the information on which they are based will often be very incomplete.

34. In handling quality changes, as in handling seasonal and discontinuous series, it is essential to show one's contributors and those who use the index numbers that an effort is being made continuously to achieve comparability in each series, and that active concern is felt about the necessity for measuring the price movement of a given base-weighted "basket" of goods—not with the price movement of some undefinable basket. We therefore consider that it is vitally important wherever price indices are developed, that careful attention should be paid to this quality problem, and that it should not be dismissed or glossed over in correspondence with trade associations or firms. A constant check must be made on the reasons for changes in price and on detailed specifications if quality changes are not to pass unnoticed. Let me conclude by giving a concrete though simple example that anyone might meet in maintaining a price index. On a given date we noticed that although the price of a certain manuscript (school exercise) book had not altered, the specification had changed from 12 to 10 pages. In consequence, after checking with the manufacturer/price contributor, we showed the price as having risen by 20 per cent.

(G) *The Presentation and Form of Publication of Price Indices*

35. It is unwise to expect to be able to adhere to any rigid plan of publication for the indices calculated. In nearly all sectors commodity indices will be useful and their publication will usually be welcomed by the trades who have co-operated in their production, but not all composite indices are equally useful. While officials may, for their own purposes, be interested in preparing an output price index for an industry (as identified in the Standard Industrial Classification) the trades concerned may consider such an index inappropriate (perhaps because of difficulties of classification). The interested trades may sometimes feel it desirable to publish only the separate indices for commodities rather than the composite index for the output of the industry group as a whole. In other

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

cases the composite index may be considered suitable for publication alongside of and in addition to the commodity indices which go to make it up. An example of the need for flexibility of approach to publication is provided by the index for plastic mouldings. The trade expressed the (acceptable) view that the index which should, strictly speaking, have included only those mouldings, etc., classified under Standard Industrial Classification trade heading 199/1, would be more meaningful and useful if items such as plastic toys (S.I.C. 193/1) and plastic toothbrushes (S.I.C. 192) were taken account of in the plastic mouldings index as well as in the indices for toys and brushes respectively. Other examples of flexibility in deciding the most suitable coverage for particular indices are provided by the broad sector output indices for "Chemical and Allied Products" (defined so as to exclude the output of oil refineries, coke ovens, etc., although these are treated as chemical and allied trades in Order IV of the Standard Industrial Classification); "Textile products other than clothing" (defined to cover the output of all textile trades in Order X except that of hosiery and other knitted goods); "Clothing and footwear" (defined to cover hosiery and other knitted goods in addition to the output of the trades in Order XII); and "Paper products" (defined to cover the output of Order XV trades with the exception of the products of printing and publishing).

36. The publication of "pairs" of input and output price indices presents a number of problems which have to be carefully borne in mind. The Government's main interest in the sphere of input/output price comparisons lies in fairly broad economic sectors, where small features of incomparability between the relevant indices tend to be unimportant. It would be extremely difficult to construct sufficiently accurate input indices for most individual trades owing to the large number of small items involved and the proportion of materials unspecified in the Census of Production reports. In the larger composites—covering broader sectors—the significant items can be more clearly defined and are less numerous, so that their construction and publication present fewer problems than the construction and publication of input indices for narrow sectors. Individual trades are also not usually in favour of publication of narrow input/output pairs, mainly on the ground that the time lag between the movement of input and output price indices is not appreciated sufficiently by those who might wish to quote them against the trade concerned. Also, in some cases input indices, particularly for food trades, may not be comparable with output indices owing to the effect of subsidies and/or duties on the one or the other, etc. By contrast, little difficulty is usually found in obtaining agreement to the publication of broad sector input indices which are found useful by most sectors of industry which, while knowing (without the help of indices) pretty well what is happening to the price of materials on their own front, often have less knowledge about what is happening to material prices in general.

37. The use of official wholesale price index numbers for contract price adjustment purposes creates some publication difficulties. These indices are now widely used in industry to settle price adjustment questions, not merely in export but also in home contracts, particularly for engineering goods and for contracts for building for local authorities where a considerable period may elapse between order and delivery. In such cases the price increase shown by the official index is now often accepted as an alternative to maintaining data on price changes for numerous individual and quite minor items. The existence of these contract price adjustment clauses and the fact that they are (without any official guidance or advice) tied to official indices, make it very important to explain and to undertake (*a*) that new indices likely to be used in contracts will not overlap with those in some old series previously used for the same purpose (thus creating alternative and different movements for indices of the same title) and (*b*) that a policy will be followed which avoids the "chopping and changing" of published figures due to revisions received of individual prices whose effect on the composite index, although very slight, might nevertheless be unsettling. The Board of Trade adopts the policy that, in order to increase the confidence of those using the main composite indices, small revisions further back than a few months should not, normally, be made. Indeed, it is very rarely that they are revised at all, although the figures may sometimes be kept in a provisional state pending firm information about certain important prices. The fact that the figures are occasionally marked as provisional is a warning to those using them for contract price adjustment purposes that they are not yet firm. Thus, although the Board of Trade always revises individual commodity indices as far as is required to take account of errors made or late notification of price changes, such revisions are only carried through to the composites if they affect comparatively recent months or if they are so important that they cannot suitably or safely be ignored.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(H) *The Ending of the Old* (1930 = 100) *Wholesale Price Series*

38. Owing to the fact that most of the old (1930 = 100) series of wholesale price indices continued to appear in the *Board of Trade Journal* for several years after the new (June 30th, 1949 = 100) series got into its stride, manufacturers found it possible, without serious inconvenience, to transfer contracts which were based on one or other of the old indices to one or other of the much more suitable indices available in the new series. Thus the termination of the old (1930 = 100)* series of wholesale price indices, when it occurred at the end of last year was not an event of any great importance.

39. Among the reasons for thinking that the "all articles" index should be discontinued were the following:

(i) A feeling that the production of any new "all articles" index would tend to give official backing to the idea that a single indicator of the general level of prices was useful, whereas it was considered that such an indicator was outdated for reasons developed previously by Mr. Stafford in paras. 12 and 13 of his May 23rd 1951 paper. With Professor Allen (see discussion of above paper) the Board of Trade considered that "a general index" was "all very well as one column on a page of United Nations statistics, along with similar indices from other countries," but did "not think that economists nowadays have much real use for an all-commodities index". Far better, it was argued, was the alternative of having separate indicators for basic materials and finished products, and for the output of and input to given sectors etc., comparing these with the parallel indicators for similar categories in other countries, which will thus be more comparable than would be the various "all articles" indices in each country. Not only are such "general" indices constructed in a completely different way in different countries, thus making international comparisons most unreal, but in many countries they exclude (as did the old index) almost all manufactured products. Thus many comparisons made with so called "all articles" or "general" indices are completely misleading.

(ii) A desire to encourage the use of the new sector indices and to discourage the over-simplification that might result from publication of a general index. The new sector indices summarize, as far as is economically meaningful, the price movements in the economy. They reveal shifts in the price structure and the dispersion or concentration of price movements for different flows of goods—of fuel or of raw materials used in industry, of different categories of output and of goods entering into various kinds of consumption. By not publishing a new "all articles" index (the construction of which would in any case be very controversial and inevitably different from that of the old index of that title), the Board encourages those who have been using the old series index to choose indices in the new series which will almost certainly prove more suitable for their purposes.

(I) *Rebasing of the Present Wholesale Price Series on* 1954 = 100

40. Finally, to conclude this part of the paper, let me say that it is the intention of the Board to recalculate all indices in the present series with a new (1954 = 100) price and weights base as soon as the detailed results of the Census of Production for 1954 are generally available. The series will then have, in place of a 1948 weights and a June 30th 1949 price base, a common price and weights base of one whole year (1954) during which prices were relatively stable after the end of the price downswing after the Korean war.

41. The question may be asked—is there a sufficiently strong case for rebasing the index at all, taking into account the considerable advantages of continuity on a single base? It is the feeling of

* For the full history of the wholesale price indices published by the Board of Trade, the following sources may be consulted: (1) *Accounts and Papers*, 1903, Vol. 33; "Report on Wholesale and Retail Prices in the U.K., 1902, with Comparative Statistical Tables for a series of years (1871–1902)". (2) *Journal of the Royal Statistical Society*, March, 1921; "The Measurement of Price Changes," by A. W. Flux, C.B. (3) *Board of Trade Journal*, January 20th, 1921, "Special Articles: Revision of Basis of Board of Trade index number." (4) *Journal of the Royal Statistical Society*, Part IV, 1933 "The Measurement of Price Changes; Retrospect and Prospect," by A. W. Flux, C.B. (5) *Board of Trade Journal*, (Supplement) January 24th, 1935, "Wholesale Prices, 1930–34; Board of Trade Index Number: a new Compilation." (6) *Journal of the Royal Statistical Society*, Part IV, 1951, "Indices of Wholesale Prices," by J. Stafford, C.B.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 1(a).—*Wholesale Price Index*

Sector Input Indices and Selected

	1949	1950		1951		1952	
	Dec.	June	Dec.	June	Dec.	June	Dec.
(a) *Sector Input Indices:*							
Basic materials (excl. fuel) used in non-food manufacturing industry	112·6	126·6	176·7	193·0	180·9	160·0	153·4
Materials used in mechanical engineering industries	105·2	111·6	125·6	131·2	142·6	151·6	148·9
Materials used in the electrical machinery industry	109·6	120·4	142·2	151·9	159·0	167·6	162·7
Materials used in the textile industries	120·0	137·4	207·2	210·1	179·5	151·5	145·4
Building and civil engineering materials	102·5	104·8	111·1	126·7	132·5	133·4	130·0
House building materials	102·7	104·0	108·5	125·3	127·9	130·5	128·3
Fuel used in manufacturing industry	..	..	..	..	..	..	123·6
(b) *Commodity Indices for Materials*:*							
Building materials:							
Bricks, fletton common, (delivered only)	100·0	106·1	106·1	117·4	117·4	122·9	126·3
,, common, other than fletton (del. only)	101·3	102·7	104·9	115·0	119·0	127·8	128·0
Cast stone and cast concrete products	100·0	98·7	101·4	104·5	107·0	113·2	115·7
Cement	96·2	96·2	104·7	122·6	128·2	130·6	122·8
Electrical porcelain	100·0	100·0	100·0	115·0	115·0	115·0	115·0
Glass, sheet and plate	100·0	100·0	97·3	98·6	108·2	111·8	120·2
Granite chippings (delivered only)	100·0	108·6	108·6	118·2	120·0	129·1	137·3
Joinery	..	..	..	123·6	125·2	123·8	115·6
Paint	100·1	103·6	112·0	131·4	138·7	136·5	135·1
Coal for carbonising	100·0	101·0	101·0	110·1	110·4	121·3	121·8
Hides and skins	103·9	116·3	176·9	176·2	122·6	95·0	115·2
Iron ore imported	..	..	..	148·9	153·9	202·1	192·5
Leather for footwear	..	..	..	187·7	146·9	112·2	127·1
Non-ferrous metals:							
Aluminium, virgin ingots, delivered	124·4	124·4	133·3	137·8	137·8	171·1	184·4
Aluminium extruded and drawn products	111·2	111·3	116·0	120·0	120·0	135·6	140·3
Copper	130·2	156·0	171·9	199·2	193·2	219·3	242·6
Copper sheets	123·9	142·6	155·7	178·9	176·1	195·0	213·8
Lead	118·3	115·6	165·9	195·1	213·4	159·4	122·1
Tin	106·7	105·8	203·4	171·3	162·3	171·0	166·8
Zinc	109·6	159·9	193·6	205·1	243·6	177·8	141·0
Brass ingots	150·6	238·7	425·7	443·4	417·5	420·3	316·0
Rubber	141·2	244·5	521·0	441·1	406·2	263·2	264·4
Sulphur, crude	100·0	114·0	114·0	146·9	202·8	190·5	190·5
Synthetic resins and plastic materials	100·3	101·6	110·8	124·4	133·4	131·6	128·0
Textiles:							
Cotton, raw	128·1	141·4	213·1	245·7	213·1	165·9	144·9
Irish linen yarns (wetspun)	105·6	116·5	145·9	186·9	152·6	..	112·4
Jute, raw	127·9	127·9	127·9	239·1	239·1	144·1	100·6
Rayon filament and staple fibres	100·0	100·0	114·7	130·5	130·5	132·1	132·1
Sisal, E. African	122·4	142·1	208·3	267·7	232·8	156·3	112·3
Wool, raw	126·7	161·9	285·3	218·0	160·1	146·8	160·9
Timber:							
Hardwood, homegrown	..	..	..	128·1	142·1	136·0	127·3
,, imported	100·0	113·6	127·3	144·6	150·8	137·0	131·7
Softwood, imported	..	..	..	..	..	151·8	151·5
Woodpulp for paper-making	66·1	78·5	137·7	221·9	226·8	164·9	102·4

* A few of these are also output price indices. They are included here since the items concerned are primarily regarded as materials, and prices to users are fairly reflected by the indices shown.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Numbers (June 30th, 1949 = 100)

Indices for Individual Materials

1953		1954		1955		1950	1951	1952	1953	1954	1955
June	Dec.	June	Dec.	June	Dec.	Annual Average	Annual Average	Annual Average	Annual Average	Annual Average	Annual Average
145·4	141·7	145·2	146·4	150·7	155·3	138·6	193·3	162·1	145·8	143·5	152·4
143·1	147·6	150·8	155·8	163·8	176·1	113·2	134·3	149·9	145·7	150·4	168·1
150·8	157·3	159·9	168·3	181·0	195·8	122·0	151·9	165·4	155·2	160·5	185·7
146·7	145·6	152·5	144·4	142·5	134·9	155·0	211·0	154·7	145·7	147·4	140·7
130·2	130·0	131·3	133·5	136·3	140·4	106·1	125·6	133·6	130·4	131·4	137·3
128·5	128·4	130·4	133·1	136·5	140·2	105·2	123·0	130·6	128·7	130·7	137·1
..	131·7	138·1	139·0	140·6	156·5	103·0	112·2	124·1	130·8	136·2	147·2
126·3	128·7	132·4	132·4	132·5	132·5	103·9	115·3	124·6	127·4	131·8	132·5
131·2	134·1	134·9	137·1	139·9	145·2	102·6	115·1	125·9	131·7	135·7	140·9
117·9	120·0	118·8	119·5	120·2	123·5	100·1	105·3	112·8	118·0	119·0	121·2
128·6	128·6	130·6	130·6	130·6	137·1	99·5	121·9	129·1	127·0	130·2	133·3
115·0	115·0	126·5	126·5	126·5	126·5	100·0	113·9	115·0	115·0	124·6	126·5
126·9	126·9	126·9	133·4	133·4	140·7	99·6	100·5	114·2	126·8	127·6	134·6
137·3	147·3	147·3	147·3	139·6	134·6	105·0	116·9	129·0	142·3	147·3	140·3
115·1	115·4	116·6	118·5	121·2	121·4	..	119·2	122·0	115·3	116·6	120·7
134·8	134·9	131·1	131·3	131·1	131·5	104·8	130·0	136·7	134·9	131·5	131·2
132·7	132·7	140·0	140·0	141·2	164·3	100·6	109·2	121·3	130·8	137·7	152·4
113·3	112·1	110·2	99·9†	96·9	99·0	129·0	167·8	102·8	112·1	106·2†	96·1
194·8	193·2	174·2	171·3	174·8	174·9	..	..	193·3	194·2	173·6	175·3
129·4	129·7	125·9	115·6	115·8	117·7	..	176·0	121·1	129·2	124·1	116·2
178·9	166·7	173·3	173·3	181·1	190·0	126·5	137·8	173·1	174·2	173·3	185·6
138·2	134·0	141·6	143·3	153·0	157·2	112·5	120·0	136·0	136·5	140·9	153·7
214·5	199·6	205·0	240·8	278·4	330·4	152·2	187·5	220·7	217·6	213·4	295·7
192·2	181·5	188·5	213·1	256·9	292·3	140·4	170·4	196·9	195·2	194·0	262·6
112·2	113·5	121·8	130·5	128·7	141·9	129·8	197·5	165·4	114·8	120·3	132·5
122·2	114·0	128·0	124·0	127·3	144·9	130·7	189·6	169·4	129·1	126·3	130·1
94·9	99·6	107·1	110·6	122·2	131·6	152·9	220·1	191·8	100·7	104·5	121·2
208·3	352·2	333·0	358·1	419·2	457·7	246·1	421·1	396·0	273·3	337·0	431·6
202·7	173·1	191·9	250·8	305·9	371·5	324·6	499·1	282·8	199·1	198·1	329·9
170·7	170·7	158·9	160·4	176·6	176·6	114·0	174·9	193·6	174·0	159·6	176·6
124·2	123·2	121·4	120·9	120·9	120·7	103·6	127·0	131·3	125·0	121·6	120·9
140·0	143·3	155·4	157·4	150·6	139·1	156·8	223·2	171·9	140·6	155·0	149·7
110·2	112·9	112·7	112·7	111·1	111·1	120·5	175·5	..	111·0	112·7	111·5
117·3	122·9	118·8	136·6	104·7	104·2	127·9	215·6	154·2	111·4	122·4	114·4
132·4	130·1	130·1	130·1	130·1	130·1	103·6	130·5	132·0	131·2	130·1	130·1
99·6	101·1	98·3	76·9	87·4	89·8	156·8	253·7	167·7	101·8	93·6	88·4
174·9	167·0	177·5	146·6	148·7	134·5	195·8	252·3	147·9	171·1	160·8	142·0
124·6	120·6	120·6	128·9	132·8	137·7	..	129·4	134·1	125·5	123·7	133·0
131·1	127·3	124·3	128·1	132·8	135·3	114·7	143·3	137·9	130·5	125·2	133·0
143·7	138·9	143·4	152·0	154·4	156·3	..	..	158·0	144·1	145·0	154·9
96·3	103·2	108·4	109·4	114·5	116·3	91·2	208·3	150·1	98·9	108·1	114·3

.. Not available. † Significant change of basis of index.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 1(b).— *Wholesale Price Index* Sector Output Indices and Selected Indices for

	1949	1950		1951		1952	
	Dec.	June	Dec.	June	Dec.	June	Dec.
(a) *Sector output Indices*:*							
All manufactured products other than fuel, food and tobacco (i)	102·0	105·5	113·4	126·5	128·4	127·8	126·3
Chemical and allied products (ii)	99·5	102·9	108·1	119·9	128·0	132·0	130·1
Iron and steel (iii)	100·2	100·7	102·1	107·5	123·8	138·9	138·4
Textile products other than clothing (iv)	109·8	123·9	163·1	190·2	160·9	134·2	129·0
Clothing and footwear (v)	99·7	102·9	109·7	128·4	126·9	112·8	111·7
Products of the food manufacturing industries (vi)				117·8	124·8	132·1	136·8
Paper products (vii)	90·7	95·7	119·2	172·0	184·7	165·4	136·6
(b) *Commodity Indices:*							
Clay Products:							
Refractory goods	100·2	102·6	102·9	111·5	117·1	128·4	128·6
China and earthenware	100·2	101·3	103·1	115·4	120·0	120·3	119·9
Chemical and allied products:							
Dyes and dyestuffs	100·0	100·0	105·0	110·7	118·1	129·9	129·9
General chemicals	101·6	110·5	112·4	125·7	131·0	138·3	138·1
Drugs and pharmaceutical preparations	94·4	94·6	94·7	103·4	105·5	104·7	105·0
Toilet preparations and perfumery	100·5	97·7	103·5	110·8	113·5	117·2	118·7
Soap	100·0	99·2	98·5	108·1	110·7	116·6	116·6
Electrical appliances and metal goods:							
Television receivers	100·0	100·0	95·7	96·3	97·5	95·1	90·1
Domestic electrical appliances	100·2	100·2	102·0	107·8	113·2	115·8	115·3
Cutlery (other than razors and razor blades)	101·0	102·7	103·5	121·3	123·7	123·7	123·7
Domestic hollow-ware	102·9	102·9	106·7	114·6	124·5	134·7	139·2
Textile products:							
Cotton cloth	..	126·3	..	194·4	176·8	134·2	122·4
Serge cloth	108·0	132·1	187·8	216·5	141·1	125·5	132·3
Rayon fabric	102·4	103·5	114·8	132·4	133·5	123·7	115·6
Linen	..	..	..	153·7	142·5	109·8	105·9
Jute piece goods	111·6	111·6	111·6	178·9	178·9	124·5	100·1
Rope, twine and net	103·8	116·8	135·9	186·9	172·5	162·2	140·9
Lace	101·6	107·0	122·6	138·6	136·0	127·6	125·0
Carpets and rugs	104·0	128·0	155·3	183·8	154·0	142·1	154·5
Narrow fabrics	104·3	108·5	128·4	158·6	154·7	144·3	135·6
Clothing and footwear:							
Made-up clothing	99·7	100·2	104·8	118·2	121·1	109·0	106·9
Hosiery and other knitted goods	98·7	107·0	117·6	148·8	133·1	110·8	111·2
Boots and shoes, etc. (other than rubber)	100·6	106·2	113·0	134·9	134·5	121·2	121·1
Rubber footwear	100·0	100·0	127·9	138·6	138·6	138·6	137·4
Timber products:							
Domestic furniture	100·0	99·8	101·9	114·0	113·9	113·2	111·2
Wooden boxes, packing cases and crates, etc.	98·8	97·9	97·0	116·5	124·3	122·5	117·5
Paper products:							
Paper and board	86·8	94·3	120·4	184·4	200·3	179·0	145·0
Manufactured stationery	96·0	100·4	119·9	164·3	166·1	156·5	134·7
Wallpaper	100·0	100·0	117·9	135·0	132·1	131·8	129·1
Cardboard boxes, cartons, fibre-boards packing cases	94·6	97·7	113·6	153·5	164·6	146·1	126·9
Rubber products:							
Rubber tyres and tubes	106·2	114·7	145·8	179·7	179·7	165·8	158·6
Rubber goods (other than tyres and tubes)	100·3	104·3	126·5	163·8	166·2	156·9	144·7
Miscellaneous products:							
Linoleum and felt base	94·6	94·6	94·6	105·6	115·6	116·3	110·4
Toys and games	..	..	..	117·1	119·8	121·8	121·4
Plastic mouldings	97·7	98·7	99·1	107·6	110·1	110·6	108·3

* The coverage of the sector output indices, in terms of Standard Industrial Classification groupings, is as follows:

(i) 20–149, 170–199 less 82 and parts of 30, 36 and 83.
(ii) 31, 32, 33, 34 and 35.
(iii) 40, 41, 43 and 44.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Numbers (June 30th, 1949 = 100)

Individual Products and Groups of Products

1953		1954		1955		1950	1951	1952	1953	1954	1955
June	Dec.	June	Dec.	June	Dec.	Annual Average	Annual Average	Annual Average	Annual Average	Annual Average	Annual Average
125·3	124·8	125·8	126·3	129·3	132·4	106·8	124·8	127·8	125·2	125·5	129·6
130·5	129·7	131·0	131·1	132·3	135·9	103·7	120·9	131·0	130·4	130·8	133·1
140·2	139·8	142·3	143·2	144·7	153·2	101·0	113·2	136·5	139·8	141·3	148·1
129·6	129·9	131·6	128·6	126·8	123·8	131·1	178·5	136·4	129·4	130·3	126·1
113·9	114·8	114·3	113·5	113·3	114·3	103·7	124·7	114·9	113·7	114·2	113·6
137·2	138·5	136·6	141·4	143·3	144·4	..	118·0	132·5	138·6	139·1	143·9
131·3	130·4	133·3	133·7	140·1	142·4	101·4	168·4	160·9	131·0	133·3	140·3
132·1	133·3	134·7	140·3	148·0	150·8	102·3	112·7	127·1	131·7	137·2	146·0
119·6	120·7	127·1	127·7	129·2	134·1	101·5	113·3	120·2	120·2	125·0	130·7
128·8	128·8	129·0	129·0	129·0	138·0	102·3	112·0	128·1	128·9	129·1	132·0
143·0	143·1	147·5	148·7	150·3	154·2	109·2	124·4	138·0	142·7	147·3	151·4
102·6	101·8	102·7	102·8	102·4	104·1	94·1	102·4	105·2	102·5	102·2	102·8
119·1	119·9	120·3	120·9	122·3	125·5	99·3	110·6	117·5	119·4	120·3	122·4
106·3	106·3	106·0	103·0	105·6	114·3	99·2	108·0	115·8	109·3	105·7	106·1
88·3	86·1	84·3	83·8	84·0	83·2	99·1	96·6	94·4	88·3	84·7	83·8
112·5	110·9	110·5	111·7	112·5	111·8	100·3	108·9	115·5	112·8	110·6	112·1
126·4	129·1	131·6	134·2	141·0	147·6	102·7	119·8	123·7	125·7	131·2	141·7
139·7	138·9	140·9	144·5	150·7	157·9	103·8	116·2	134·7	139·3	141·3	151·6
117·3	121·2	124·7	126·5	123·1	120·2	..	..	136·9	119·1	125·0	122·8
145·3	141·1	140·7	130·1	128·8	121·4	140·5	196·4	128·5	141·3	137·4	126·7
115·1	117·0	117·1	113·4	112·8	112·0	105·1	132·1	121·6	115·7	116·1	112·7
101·5	103·0	102·4	102·6	101·4	99·8	..	145·0	115·1	102·6	102·4	101·3
108·0	111·2	109·7	119·8	104·8	104·5	111·6	164·4	130·5	105·5	111·7	108·9
122·5	122·7	122·4	121·6	121·0	123·1	118·4	171·1	158·4	127·0	122·5	122·2
119·1	117·7	116·8	117·1	116·6	113·1	108·5	134·2	129·0	119·3	116·8	115·2
155·0	162·8	161·5	161·5	161·5	168·0	130·0	170·2	147·3	157·7	162·0	163·6
127·8	127·7	128·6	129·5	131·0	131·9	112·1	152·5	142·8	129·2	128·6	131·5
109·5	110·1	110·1	109·4	109·4	110·3	101·0	116·6	110·6	109·1	109·9	109·7
114·8	116·2	114·6	113·2	111·6	111·7	107·5	138·7	114·1	114·5	114·8	111·9
122·1	123·2	122·9	122·2	122·4	123·6	106·6	131·5	123·8	122·5	122·8	122·8
130·0	129·2	125·6	125·6	131·3	140·7	105·2	136·4	138·3	132·3	126·9	132·2
108·4	107·1	104·9	105·3	106·2	109·4	100·3	111·2	112·5	108·8	105·3	106·8
112·7	110·7	108·6	109·4	113·0	118·3	97·9	116·6	121·7	113·1	109·2	113·5
141·2	139·5	141·7	141·9	147·8	150·9	100·4	181·2	173·7	140·7	141·8	148·4
127·4	124·8	128·9	128·9	137·1	138·1	106·0	158·7	151·6	126·9	128·2	136·5
129·1	117·5	117·5	117·5	127·7	131·6	103·8	126·6	131·4	125·2	117·5	128·7
120·7	125·1	128·0	129·6	135·1	137·4	100·5	150·1	143·8	121·9	128·2	135·1
138·4	138·4	138·4	145·3	156·5	168·3	123·0	179·7	168·8	141·9	139·8	161·3
136·8	134·8	134·0	134·7	142·5	151·9	108·6	159·7	155·6	138·0	134·4	144·0
103·2	103·2	98·2	98·2	98·2	100·6	94·6	106·2	115·5	103·7	98·8	99·3
115·9	116·2	117·4	117·4	121·2	124·7	..	116·2	121·2	116·3	117·0	121·8
108·1	106·6	106·8	107·4	109·0	109·6	98·9	107·4	110·5	108·0	106·8	108·6

(iv) 110–117, 119–129.
(v) 118, 140, 141, 142, 147, 148 and part of 190/2.
(vi) 150 to 162, inclusive.
(vii) 180, 181, 182 and 183.

.. Not available

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the Board that the years 1948/1949 were a little too near to the end of the war and a little too early in the period of comparatively rapid industrial change following it, to be satisfactory as a base for indices which could survive for a long period of years. The change to a 1954 base should coincide with a period of much greater stability for the index, and the arguments for subsequent changes in the weighting base may be weaker. Accordingly, although it is not expected that the changes in weights involved in transferring to a 1954 = 100 base will be at all radical (except on a few sectors), it is felt both from the point of view of the structure of the indices and their movement that this rebasing is desirable.

PART II

THE RELATIVE MOVEMENTS OF THE MAIN WHOLESALE PRICE INDEX NUMBERS IN THE (JUNE 30TH, 1949 = 100) SERIES WITH SOME ANALYSES AND COMPARISONS

42. This part of my paper is concerned with setting out and examining the new series of wholesale price indices themselves and extracting from them various facts of statistical and economic interest.

(1) *A Selection of the Published Index Numbers*, 1949–1955

43. The first step towards an appreciation of the new series of wholesale price indices seems to be that of summarizing the movements of the main indices since their inception. It may be convenient and useful to bring together and to set these out, for June and December, together with the annual averages, for each year since 1949. This is done at Tables 1(*a*) and 1(*b*). These tables show sector input and output indices as well as indices for individual materials and products.

(2) *Simplifying the Construction of Major Composite Index Numbers*

44. Having constructed indices on a tolerably thorough basis and having found that in some cases one would have obtained almost the same index by using fewer indicators, it is not unnatural to wonder whether major indices could not, generally, be safely constructed from fewer price series than is usually the case. Examination of the indices makes it clear that it is certainly possible to find very good methods of constructing major price indices which may, for some purposes, prove useful and time conserving, but that this is only possible *after* they have been constructed on a thorough basis. To try to streamline these indices before calculating them properly might well produce very false answers. There is no doubt that the large number of series for individual commodities and products available as a result of a thorough process of construction is extremely useful in its own right, apart from increasing the accuracy of the major series, but it is certainly true that sound indicators (for most practical purposes) can be constructed by using only a selection of the major items that feature in the leading composite indices. The wise choice of these indicators to some extent presumes experience of the movements of all the indicators included in the index, however, and it should not be assumed that, simply because there are very small differences between the movements of an index constructed in one way or another, these small movements may not be significant. In particular periods (e.g. when the index is comparatively steady) or for particular purposes, these small differences may be significant. While it is sometimes possible for major indices to be simplified so far as the number of items included in them is concerned, it is also often possible for *slight* differences in the weighting of particular indices, to allow for duplication, or differences in coverage, to be ignored with no serious practical effect. There comes a point, however, where the consequences of further streamlining become much more serious and it would be very difficult to know when this point was likely to be reached unless one first experimented with the full range of data. The examples in Table 2 may help to illustrate some of the points made above; in particular, columns 3, 4 and 6 help to show what substantial differences too radical methods of simplification can make. Despite such examples, it is clear that too much importance should not be attached to the minor series which may be used in constructing sector indices.

TABLE 2

(3) *A Comparison of the Price Movements of Related Groups of Increasingly Processed Commodities*

45. (*a*) The fact that the price movements of finished goods are often "damped" compared with those of less processed items of the same "kin" or with the primary materials entering into them, is widely appreciated. The figures in Table 3, relating to the relative levels of certain indices, illustrate

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 2

A Comparison of Selected Major Wholesale Price Indices when Calculated in Different Ways

Selected Months	All Manufactured Products (other than fuel, food and tobacco)		Basic Materials (excluding fuel) for non-food and tobacco manufacturing industries				Textile Materials		Textile Products (i)		Iron and Steel (j)	
	Full (published) index (a) (June 30th, 1949=100)	Index calculated from selected indicators (b) (June 30th, 1949=100)	Full (published) index (c) (June 30th, 1949=100)	Index calculated from 34 indicators only (d) (June 30th, 1949=100)	Index calculated from 10 indicators only (e) (June 30th, 1949=100)	Old (1930=100) sub-index for "Basic Materials" (f) (Converted to June, 1949=100)	Published index for materials used in the textile industries (g) (June 30th, 1949=100)	Basic textile materials sub-group of basic materials Index (h) (June 30th, 1949=100)	Published index for textile products other than clothing (June 30th, 1949=100)	Textile products sub-group of all manufactured products index (June 30th, 1949=100)	Published index for iron and steel (S.I.C. headings 40, 41, 43 and 44) (June 30th, 1949=100)	Combination of separate published figures for S.I.C. headings 40/41, 43/1, 43/2 and 44 (June 30th, 1949=100)
1	2	3	4	5	6	7	8	9	10	11	12	13
Dec., 1949	102·0	101·8	112·6	112·7	114·2	103·7	120·0	125·6	109·8	109·9	100·2	100·3
June, 1950	105·5	105·9	126·6	127·5	132·5	122·1	137·4	146·8	123·9	123·4	100·7	100·6
Dec., 1950	113·4	112·9	176·7	178·2	197·2	191·5	207·2	235·1	163·1	160·5	102·1	101·9
June, 1951	126·5	127·4	193·0	191·2	204·0	209·5	210·1	230·9	190·2	187·8	107·5	107·0
Dec., 1951	128·4	127·4	180·9	178·1	180·1	189·7	179·5	188·8	160·9	161·8	123·8	124·2
June, 1952	127·8	127·2	160·0	157·5	159·5	149·7	151·5	153·6	134·2	133·7	138·9	139·7
Dec., 1952	126·3	125·9	153·4	151·6	156·0	136·0	145·4	148·3	129·0	127·8	138·4	139·2
June, 1953	125·3	125·7	145·4	144·3	151·1	126·7	146·7	152·0	129·6	127·7	140·2	141·1
Dec., 1953	124·8	125·4	141·7	141·2	149·2	125·8	145·6	150·4	129·9	128·4	139·8	140·7
June, 1954	125·8	126·4	145·2	145·6	155·7	127·9	152·5	158·7	131·6	129·9	142·3	142·9
Dec., 1954	126·3	127·5	146·4	146·4	155·3	127·8	144·4	146·9	128·6	127·5	143·2	143·7

(a) Constructed from several hundred price indices based on several thousand prices.

(b) Ignoring about 100 indices carrying approximately one-third of the total weight in the full index.

(c) Constructed from 76 price indices some of which, e.g. those for wool, cotton and home-grown hardwood, are the combination of a number of individual prices series.

(d) Selected from the 76 price indices making up the full published index; the 34 indices selected carry 90 per cent. of the total weight in the full index.

(e) Selected from the 76 price indices making up the full published index; the 10 indices selected carry approximately 62 per cent. of the total weight in the full index.

(f) This sub-index of the old (1930 = 100) index using the old weights was published regularly up to and including April, 1951. Nineteen of the twenty-four price indices making up this index also feature in the new (June 30th, 1949 = 100) Basic Materials index where they carry a weight of 52·6 per cent. Their weights relative to one another differ radically, however, in the two indices, e.g. wool and cotton items, carried weights of 15·2 and 6·1 per cent. respectively in this index, and 14·4 and 12·2 per cent. in the new series index now published.

(g) Includes important non-textile materials such as dyes and dyestuffs, chemicals, wood-pulp and packing materials.

(h) The series featuring in this index are also included in the index for "Materials used in the Textile (other than clothing) industries", where they carry a combined weight of approximately 75 per cent., the balance being mainly the "non-textile" materials referred to above (g).

(i) These two indices for textile products differ to the extent that whereas the weighting of the published index for "Textile products other than clothing" excludes the value of products transferred within the textile industries sector only, the weighting of the textiles sub-group of the "All Manufactured Products" index excludes the value of products transferred within manufacturing industry as a whole. Thus, for example, the value of cotton cloth passing to the clothing industry is included in the former index but excluded from the latter.

The comparative 1948 weights of the three most important products (each carrying a weight of over 10 per cent. in both indices) are as follows:

Textile products other than clothing index:

Cotton cloth 27·2 per cent.; Wool cloth 18·8 per cent.; Rayon fabric 11·3 per cent.

Textile products sub-group of All Manufactured Products index:

Cotton cloth 30·9 per cent.; Wool cloth 14·0 per cent.; Rayon fabric 10·9 per cent.

The lower proportion of wool cloth in the textile sub-group of the All Manufactures Products Index reflects the large amount of this cloth entering the clothing industry in 1948. The value of this cloth is excluded from the All Manufactured Products but included in the Textile Products Index as it passes out from the textile to the clothing sector. The proportion of cotton cloth going into the clothing industry, though substantial, was smaller than wool cloth in 1948.

(j) The same group weights—40/41 (63·4 per cent.), 43/1 (13·3 per cent.), 43/2 (7·9 per cent.) and 44 (15·4 per cent.)—have been used in calculating both these indices. The composition of 40/41 differs, however, in the two indices, mainly to the extent that in the index published separately for this heading ("Products of blast furnaces and iron and steel melting and rolling") tinplate and sheet bars are included with a weight of 13·5 per cent. These items are not represented in the published composite index for "Iron and Steel" since they are further processed within the iron and steel sector and are accounted for in the value of the tinplate and sheets which they produce.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

TABLE 3

*Comparison of Price Movements of Related Groups of Increasingly Processed Commodities**

Summary Analysis for Selected Indices

(1) Commodity Groups	(2)* Level of indices at "Peak" taking peak level of index for primary material in each group as 100 (a)	(3) Lag in months, if any, between month of peak index number for specified item and month of peak index number for primary material (negative lag shown in brackets)	(4)* Level of indices at subsequent low point, taking low point of index for primary material in each group as 100	(5) Lag in months, if any, between subsequent low-point of index for specified item and month of low-point index number for primary material (negative lag shown in brackets)	(6)* Annual averages of indices in main year of (Korean) price rise (1951) taking 1951 annual average of index for primary material in each group as 100	(7)* Annual average of indices in year following end of Korean price fall (1954) taking 1954 annual average of index for primary material in each group as 100
Aluminium:						
1. Aluminium ingots	100	—	100	—	100	100
2. Aluminium extrusions	78	1	80	0	87	81
3. Aluminium forgings	70	0	75	1	82	74
Cotton:						
1. Raw cotton	100	—	100	—	100	100
2. Cotton yarn	89	1	89	0	92	87
3. Cotton cloth	78	1	85	0	81	81
4. Cotton hosiery	55	3	(83)	(4)	58	74
5. Cotton made-up clothing	53	7	(80)	(4)	55	72
Wool:						
1. Raw wool	100	—	100	—	100	100
2. Worsted yarns	82	0	99	0	91	96
3. Serge cloth	62	1	96	0	78	85
4. Wool hosiery	42	3	88	4	62	78
5. Wool made-up clothing	30	6	84	5	46	70
Jute:						
1. Raw jute	100	—	100	—	100	100
2. Jute piece goods	75	0	102	0	76	91
3. Canvas goods and sacks	73	6	110	9	75	88
Textiles:						
1. Materials used in the textile industries	100	—	100	—	100	100
2. Products of the textile industries, other than clothing	74	1	89	0	85	88
3. All clothing (other than footwear)	48	5	(75)	(5)	58	75
Leather:						

3. All clothing (other than footwear) . .	48	5	(75)	(5)	58	75
Leather:						
1. Hides and skins . .	100	—	100	—	100	100 (b)
2. Leather for footwear .	90	1	131	0	105	117
3. Travel goods . .	99	1	147	4	78	117
4. Other leather goods .	53	2	131	3	70	103
5. Boots and shoes .	61	5	144	5	78	116
Timber:						
1. Imported and home grown timber . .	100	—	100	—	100	100
2. Joinery . . .	80	2	85	0	84	84
3. Domestic furniture .	(72)	(4)	78	7	79	76
Paper:						
1. Woodpulp . .	100	—	100	—	100	100 (b)
2. Paper and board .	(87)	(3)	146	5	87	131
3. Manufactured stationery . . .	(73)	(6)	130	5	76	118
Rubber:						
1. Raw rubber . .	100	—	100	—	100	100
2. Rubber tyres and tubes	(27)	(1)	(85)	(8)	36	71
3. Rubber goods other than tyres and tubes	25	7	82	8	32	68
4. Rubber footwear .	21	3	77	4	27	64
Plastics:						
1. Plastic materials .	100	—	100	—	100	100
2. Plastic mouldings .	83	1	(88)	(11)	85	88
All Products Non-food Manufacturing Industry:						
1. Basic materials . .	100	—	100	—	100	100
2. Manufactured products	59	12	89	0	65	88

* It should be emphasised that the indices shown in columns 2, 4, 6 and 7, show the relationship of the levels of related *indices*—and have nothing to do with the relative levels of absolute *prices*.

(a) Brackets denote peak or low-point index numbers for processed commodities which occur before to those of their related raw materials. This negative lagging in some cases reflects no more than a long period averaging out of anticipated costs (possibly due to controls, in a trade, but it can also reflect the fact that the quantity of materials bought at peak prices (especially in the more volatile cases) was relatively low. It will be seen that rubber and paper products, as well as furniture, are involved.

(b) In both these cases the price indices for the primary material (hides and skins and woodpulp respectively) fell to a relatively low level in the post-Korean period. Price indices for the more processed commodities, embodying as they do a high proportion of labour costs, naturally (in a period of rising wages) fell less violently. Consequently, the indices for these products stood at a higher level in 1954 than those for their respective raw materials.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar
Digitized by Arya Samaj Foundation Chennai and eGangotri

Digitized by Arya Samaj Foundation Chennai and eGangotri

this. Not merely are violent "rises" in the prices of raw materials normally reflected in less violent increases in the prices of semi-manufactures and much less substantial increases in the prices of finished goods using such materials, but price "reductions" in finished products are also normally much less marked than in the materials used in their manufacture. These "damped" price movements reflect not only the effect of factors such as the productivity of labour and of management, or the scale of production, on the price of manufactured articles, but also the effect of wage movements. The cost of labour plays a large role in helping, with other factors, to determine the price of processed commodities and these labour costs are often much less volatile than those of materials. Thus, although material prices might fall very violently, after rising, the prices of finished products would only fall slowly, or not at all if wages were rising. An example of this may be seen in the relative trends of the hides and skins and footwear indices. It is not, however, suggested that there is any general "damping" rule. If, for example, material prices (excluding food and fuel) are almost stationary over a longish period, during which wage rates are rising or falling—or productivity is changing significantly—then the prices of the more processed commodities could well move more substantially than those of less processed commodities or of materials. The movement of final food prices in 1955 may be a case where—in contrast to the trend in other sectors—the prices of final consumption goods moved more substantially than those, on average, of the materials used in producing them.

(*b*) The question of time lags is not one which I propose to develop in this paper, but Table 3 makes it clear that lags of 3 to 6 months between the price movements of raw materials and their related finished products, are quite common. The time lags referred to depend, in each trade, on a number of factors, among them being the structure and final product pricing policy of that trade. This latter policy will clearly be different in trades where raw material prices are very volatile from those where the prices of materials are more stable. Changes in the prices of finished products which anticipate changes in the prices of the related materials (illustrated by the negative lags shown in brackets in Table 3 and more common when prices were falling than when they were rising) can usually be attributed to some special feature of the trades concerned. A detailed examination of such cases cannot be attempted here, but the long-term pricing policy and other factors involved will be readily detected by those acquainted with the trades where this anticipation has at times occurred.

(*c*) In further development of the answer to this question about the relationship of the levels of different indices at given dates, it is interesting to consider the relative levels of the price movements of (i) basic (non-food) manufacturing materials, (ii) semi-finished products and (iii) all (non-food) manufactured products. With the Korean peak of the basic materials index (220·4 at March, 1951) taken as 100, it is found that the average Korean peak of selected semi-finished product indices was approximately 68 and the Korean peak of the index for all manufactured products was approximately 59 (at March 1952). Similarly, if the 1954 annual average of the basic materials index (143·5) be taken as 100, then the annual average for 1954 of the same selected semi-finished products was 97 and that of the index for all manufactured products approximately 87. Few better periods can be found for the post-Korean comparison than the annual average for 1954 since, despite the fact that there was a lag of 12 months between the peak price for raw materials and that for manufactured products, taking these groups as a whole, there was no difference between the date (February, 1954) at which the basic materials index and the manufactured products index reached their post-Korean low levels. The relationship in the latter period emphasizes the smaller differences between the indices as prices regained more normal levels after the Korean war.

(4) *The Relative Volatility of the Prices of Selected Commodities*

46. (*a*) What constitutes price volatility in a commodity is by no means self-evident and the information provided at Table 4, on page 262 is intended to do no more than assist in providing an answer. When commodities are sufficiently erratic and changeable in price behaviour to merit the description "volatile", this must spring from the unpredictable and changeable nature of the supply of and demand for them. Thus, it is not surprising that the only three major commodities whose price movements, in the five and a half years ending December, 1954, have changed direction four or more times, *each time by more than* 20 *per cent.*, are raw rubber, raw wool and hides and skins. These are commodities whose prices, like the demand for and supply of them, are

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

variable—*par excellance*. In order to bring as many as ten of the commodities shown in Table 4 into this particular type of volatility category we have to lower the limits of variation taken as our test to as low as 5 per cent. (on each of four occasions between June, 1949 and December, 1954). By thus lowering our limits we bring into the specially volatile category, brass ingots, lead, raw silk, worsted yarn, shellac and tin ingots, and—if 1955 is also taken into account—copper also joins the list.

(*b*) To apply the particular measure referred to above too enthusiastically, however, especially over the five and a half years ending December, 1954, with a view to determining the relative volatility of different commodities, would be misleading if only because price controls up to (and beyond) 1952 damped the natural price variability of many commodities. For short periods and for minor causes, price controls often smoothed out price movements, though as a result the changes, when they did occur, were often more substantial than they might otherwise have been. For this and other reasons a measure, described as "A" at Table 4 (see footnote (1) to that table) on page 262, has been used and this places emphasis on that aspect of the price volatility of each commodity which is concerned with the extent of the fall between its highest and lowest levels, rather than on the number of times it has changed, in a given period.

(*c*) While it is true that the great majority of primary commodities that show considerable price volatility in response to large supply and demand influences, such as strikes, rumours of wars and wars, slumps, etc., also respond to small influences, it may, of course, be true that the price volatility of some commodities is not really tested by small influences—i.e. that a commodity the price of which shows surprising volatility in a war situation may show itself extremely stable in more normal situations, thus normally qualifying for consideration as a non-volatile commodity. As one cannot generalize about volatility in absolute terms—and as the post-1949 periods do afford an exceptionally good opportunity for studying the effect on the price of commodities of a substantial but, nevertheless, not castastrophic upheaval (prices were not comprehensively controlled during the Korean war as they were during the World War)—the selection of measure "A" does seem to be the most useful for this period. Other—equally valid—simple measures, merely emphasize the many-sidedness of price volatility.

(*d*) Comparison of the price volatility of "materials", "intermediate products" and "finished products" as groups, is not perhaps very fruitful, as the price volatility of each of these commodity categories is essentially different. Table 4 does at least confirm clearly, however, that the sort of degree of volatility that one can expect in basic materials is substantially damped in the case of intermediate products and still further damped in the case of finished products. If any generalizations were justified by the figures as they stand at Table 4 they would surely be that (i) semi-finished products and manufactures derived from volatile materials are more than usually volatile (especially textile, rubber and paper products), but that (ii) most manufactured products are either only slightly volatile or essentially non-volatile. In all cases the movement of indices has been taken to be a fair measure of the price movement of the commodities included in them. This is, of course, less than precise for composite groups of materials or manufactured products, since the prices of one or more of the many items represented in any such index might have shown a much greater or much smaller volatility than the weighted average for all the items in the index taken together. My experience suggests that this possibility is a very marginal one in the case of manufactured products—but a real one in the case of materials. In short, to complete this exercise satisfactorily one would need to consider the relative volatility not merely of, for example, "Imported Softwood" and "Imported Hardwood" but of, for example, "Douglas Fir", "Finnish Birch", "African Mahogany" and "Slavonian Oak", or possibly even specific grades of each.

(5) *Comparisons Between Wholesale and Retail Price Indices*

47. Only a limited number of useful comparisons can be made between official wholesale and retail price indices in the United Kingdom. Not only is there some significant difference in scope between indices with comparable titles, but the prices used in the retail indices necessarily include wholesale and retail margins, purchase tax, etc., all of which are excluded from the manufacturers' prices used in constructing the wholesale price index numbers. Nevertheless, two quite useful comparisons can be made. Chart 1 compares the as yet unpublished wholesale price index for Manufactured Personal Consumer Goods (other than fuel, food, drink and tobacco) with an *ad*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 4.—*The Relative* Price Volatility of Selected Commodities and*

BASIC MATERIALS					INTERMEDIATE	
	Selected measure (A) of price volatility		Grouping by alternative measures (B) and (C) of price volatility			Selected measure (A) of price volatility
Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	(A) % fall (i) between highest and *subsequent* lowest price index number	Price volatility grouping by measure (A)	(B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	(C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954	Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	(A) % fall (i) between highest and *subsequent* lowest price index number
1	2	3	4	5	6	7
Raw rubber	75	I	I	I		
Raw wool	70	I	I	I		
Esparto	70	I	II	III		
Hides and skins	63	I	I	I	Worsted yarns	63
Zinc	62	I	II	II		
Whale oil	62	I	V	V		
Linseed oil	62	I	III	IV		
Raw jute	60	I	IV	IV		
Shellac	60	I	I	I		
Group I	(60–75)					
Woodpulp	58	II	II	II		
East African sisal	56	II	II	III		
Molasses	53	II	VI	VI	Brass ingots	53
					Serge cloth	53
					Jute yarn	52
					Coir yarn	50
Lead	49	II	II	II		
					Leather for footwear	47
					Jute piece goods	46
Raw cotton	45	II	I	III	Cotton yarn	45
Group II	(45–59)					
					Irish linen yarn	41
Tin ingots	40	III	I	II	Cotton cloth	40
					Ethyl alcohol (industrial)	38
Palm oil	37	III	IV	IV		
					Linen	34
Raw silk	33	III	I	I		
Group III	(30–44)				Paper and board	30
					Pressed wool felt	28
Sulphur	22	IV	VI	VI		
Copper	20	IV	III	III		
Group IV	(20–29)					

* For simplicity the phrase "Price Volatility" is used throughout this Table, but it should be emphasized (see Paragraph 46 (*d*)) that the volatility measured is in all cases that revealed by the price *indices* for the materials and products listed, some of which—being weighted average price movements—may hide the differing volatility characteristics of their component commodity series.

Note.—(i) For the great majority of commodities the interpretation of measure "A" is simple, since there was a definite "peak" price (generally during the Korean war period) and a clear "subsequent" lowest price. In some cases, however, discretion has to be exercised in deciding what point represents the subsequent low point to which the particular commodity index fell, since in a certain number of cases prices of commodities tended to fall in several "steps". When, in such cases, it is quite clear that any downward "step" was only an "incident" on the further downward movement of the price it has been ignored. In short, the lowest price reached after the major rise in 1950/51/52 and *before the commodity rose in price* significantly for a second time is taken as the "subsequent" lowest price index. Where this subsequent fall is followed by a significant rise in price which continues for two or more months, further falls beyond this rise are not taken into account. Where an index has not fallen during the period since June, 1949, but has risen consistently, or has risen in steps consistently, then it follows that there is no lowest "subsequent" index number and that these indices must be classified as having

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Groups of Commodities between June 1949 and December 1954

PRODUCTS			FINISHED PRODUCTS				
Selected measure (A) of price volatility	Grouping by alternative measures (B) and (C) of price volatility			Selected measure (A) of price volatility		Grouping by alternative measures (B) and (C) of price volatility	
Price volatility grouping by measure (A)	(B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	(C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954	Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	(A) % fall (i) between highest and *subsequent* lowest price index number	Price volatility grouping by measure (A)	(B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	(C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954
8	9	10	11	12	13	14	15
I	I	III					
II	I	III					
II	I	III					
II	IV	IV					
II	I	I					
II	I	II					
II	IV	IV					
II	I	IV					
III	II	V	Canvas goods and sacks	41	III	II	III
III	I	III					
III	VI	VI					
			Textile products (other clothing)	35	III	I	III
III	II	III	Rope, twine and net	34	III	I	III
III	II	V					
IV	IV	IV					
			Manufactured stationery	25	IV	III	IV
			Carpets and rugs	24	IV	III	V
			Rubber tyres and tubes	23	IV	IV	V
			Brushes and brooms	21	IV	II	III
			Other rubber goods	20	IV	II	III

a difference of "0" per cent. This will be seen to be true of only a few basic materials, but of several intermediate and a large number of finished products. The groupings selected are shown at Col. 2.

(ii) The groupings selected here were as follows: Group I 0–5 months unchanged; Group II 6–20 months unchanged; Group III 21–35 months unchanged; Group IV 36–45 months unchanged; Group V 46–55 months unchanged; Group VI 56–60 months unchanged; Group VII 61 months or more unchanged. For purposes of defining "months unchanged", the first month at any particular level does not count. Thus if the index for any commodity stayed at the same level for six consecutive months in a given year it would count as having remained unchanged for "five" months in that year, and so on for each new level at which it might settle temporarily.

(iii) The groupings used here were as follows: Group I 20–30 changes of direction; Group II 15–19 changes of direction; Group III 10–14 changes of direction; Group IV 5–9 changes of direction; Group V 3 or 4 changes of direction; Group VI 1 or 2 changes of direction; Group VII no changes of direction. The phrase "change of direction" has been interpreted as meaning a change from "rising" to "falling" in price, or vice versa; it has not included a change from rising or falling to "remaining steady". Where commodities have risen in price consistently they have been classified to Group VII, even though their prices may have shown a number of months unchanged at any stage in their upward course.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 4 (continued).—*The Relative* Price Volatility of Selected Commodities and*

Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	Basic Materials: Selected measure (A) of price volatility: (A) % fall (i) between highest and *subsequent* lowest price index number	Basic Materials: Selected measure (A) of price volatility: Price volatility grouping by measure (A)	Basic Materials: Grouping by alternative measures (B) and (C) of price volatility: (B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	Basic Materials: Grouping by alternative measures (B) and (C) of price volatility: (C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954	Intermediate: Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	Intermediate: Selected measure (A) of price volatility: (A) % fall (i) between highest and *subsequent* lowest price index number
1	2	3	4	5	6	7
Imported hardwood	19	V	I	II		
Imported softwood	18	V	III	IV	Spun rayon yarn	18
Pyrites	17	V	V	IV		
Imported iron ore	16	V	I	I		
Home-grown hardwood	15	V	IV	II	Rayon fabric	15
					Wooden boxes and crates	14
Group V	(11–19)				Joinery	11
Aluminium ingots	10	VI	V	VI	Synthetic resins and plastic materials	10
Imported plywood	8	VI	IV	V	Cement	8
Carbon black	7	VI	V	V		
					Paint for building and decorating	6
Group VI	(5–10)					
					Sulphuric acid	4
					Iron and steel (tinplate)	4
					Iron and steel (sheets, plates and castings)	2
					Rayon filament and staple fibre	2
Nickel	0	VII	V	VII	Dyes and Dyestuffs	1
China clay	0	VII	VI	VII	General chemicals	0
Lime, chalk lump	0	VII	V	VII		
A *few* others	0	VII	..	..	*Several* others	0
Group VII	(0–4)					

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Groups of Commodities between June 1949 and December 1954

PRODUCTS			FINISHED PRODUCTS				
Selected measure (A) of price volatility	Grouping by alternative measures (B) and (C) of price volatility			Selected measure (A) of price volatility		Grouping by alternative measures (B) and (C) of price volatility	
Price volatility grouping by measure (A)	(B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	(C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954	Commodity or group of commodities shown by groups in descending order of volatility as determined by measure (A)	(A) % fall (i) between highest and *subsequent* lowest price index number	Price volatility grouping by measure (A)	(B) Grouping by number of months in which index remained unchanged (ii) between June, 1949, and December, 1954	(C) Grouping by number of changes of direction (iii) of index between June, 1949, and December, 1954
8	9	10	11	12	13	14	15
V	III	IV	Soft furnishings	18	V	II	IV
			Lace	17	V	II	IV
			Clothing, other than footwear	16	V	II	IV
			Linoleum and felt base	16	V	V	V
V	II	IV	Spectacles	15	V	VI	VI
V	I	II					
			Boots and shoes	13	V	II	IV
			Soap	12	V	IV	VI
V	II	IV					
VI	II	III					
			Leather goods	9	VI	II	IV
			Rubber footwear	9	VI	V	VI
VI	IV	V	Domestic furniture	8	VI	II	III
VI	II	II					
			Domestic electrical appliances	5	VI	II	III
			Toys and games	5	VI	III	III
			Plastic goods	5	VI	II	III
			Spoons and forks	5	VI	III	VI
VII	VI	VI	Drugs and pharmaceutical preparations	4	VII	III	I
VII	IV	III					
VII	III	IV	Disinfectants	3	VII	V	V
VII	VI	VI					
VII	VI	V					
VII	II	III	Cutlery	0	VII	IV	VII
			Glass	0	VII	III	IV
			Polishes	0	VII	V	VII
			Toilet preparations and perfumery	0	VII	III	IV
			China and earthenware	0	VII	IV	V
			Domestic hollowware	0	VII	II	IV
VII	..	..	Matches,	0	VII	III	II
			A *large number* of others	0	VII	..	..

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

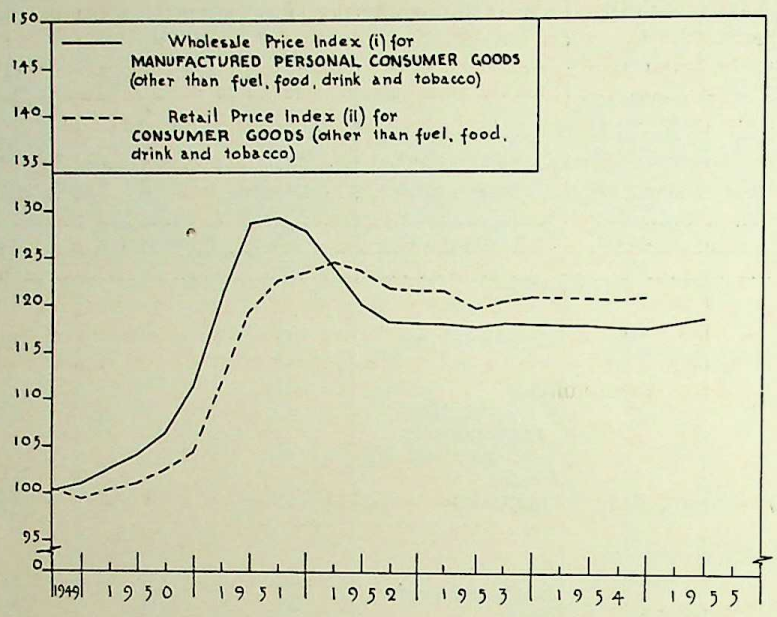


CHART 1.—(i) Board of Trade Index (June 30th, 1949 = 100). (ii) Calculated as an *ad hoc* index by combining three of the published sub-indices of the Ministry of Labour Interim Index of Retail Prices, namely those for clothing, household durables and miscellaneous goods, and converting to base June, 1949 = 100.

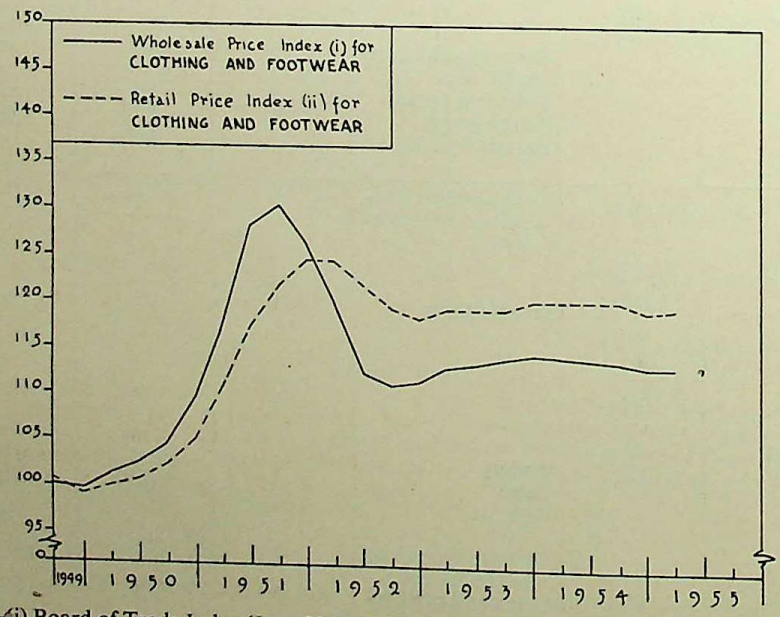


CHART 2.—(i) Board of Trade Index (June 30th, 1949 = 100). This index excludes clothing material. (ii) "Clothing" sub-index of Ministry of Labour Interim Index of Retail Prices, converted to June, 1949 = 100. In addition to footwear, this sub-index includes clothing materials with a weight of 6 in 98.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

hoc index which combines appropriate sub-groups of the Ministry of Labour Interim Index of Retail Prices; Chart 2 compares the retail and wholesale prices indices for Clothing and Footwear. Some of the figures on which these two charts are based are set out at Table 5, but it must be emphasized that the retail figures are intended to be illustrative only, as the linking of the old (June 17th, 1947 = 100) and new (January 15th, 1952 = 100) series of retail prices indices is necessarily crude and the resultant figures do not constitute official indices in any way.

48. In both comparisons, as might be expected, the extra margins (accompanied by other factors) have had the effect of damping the movement of the retail price indices. The time lag between movements of manufacturers' and retail prices varies from trade to trade and time to time (being different in some trades when prices are rising from when they are falling), but is not uncommonly about three to five months, a lag which is apparent for certain periods in the case of the Clothing indices compared at Chart 2. The fact that there are changes in time lags, margins, delivery costs, purchase taxes, etc , has to be kept in mind in attempting to forecast future retail price movements from known wholesale price movements, but it does not normally prevent approximate forecasts of this kind being made successfully.

TABLE 5

Comparison (at Six-monthly Intervals) of the Movement of Wholesale and Retail Prices Indices (1949–1954)

Date	Wholesale Price Index for "Manufactured Personal Consumer Goods"*	Retail Price Index for selected groups of items (i)	Clothing and Footwear	
			Wholesale Price Index (ii)	Retail Price Index (ii) (v)
	(June 30th, 1949=100)	(Converted from June 17th, 1947=100 to June, 1949=100)	(June 30th, 1949=100)	(Converted from 17th June, 1947=100 to June, 1949=100)
Dec. 1949	101·0	99·5	99·7	99·0
June 1950	104·0	101·1	102·9	100·6
Dec. 1950	111·7	105·3	109·7	105·1
June 1951	128·8	119·1	128·4	117·2
Dec. 1951	128·0	124·3	126·9	124·5 (iii)
June 1952	120·0	123·7	112·8	121·7
Dec. 1952	118·0	121·3	111·7	118·5 (iv)
June 1953	117·7	119·8	113·9	118·9
Dec. 1953	118·0	120·7	114·8	119·7
June 1954	117·8	120·4	114·3	119·7
Dec. 1954	117·5	119·9	113·5	119·1

* This index has not yet been published in the *Board of Trade Journal*.

(i) See footnotes to Chart 1 on p. 266.
(ii) See footnotes to Chart 2 on p. 266.
(iii) The Korean peak of this retail price sub-index was reached in December, 1951. The peak of the wholesale prices index was reached in August, 1951, suggesting a "lag", at that time, of 4 months.
(iv) The subsequent low-point of this retail price sub-index was reached in January, 1953, the low-point of the wholesale prices index being reached in September, 1952, again four months previously.
(v) The effect on the retail price index for clothing of the introduction of the D level tax, after the ending of the Utility Scheme in March, 1952, was negligible.

(6) *A Comparison of Various Published Price Indices for Raw Materials*

49. (*a*) The differences between the Board of Trade wholesale price index for "basic materials" and indices such as those published by *The Statist* and *The Financial Times*, or by the *Reuter* and *Moody* agencies, are not always appreciated and it may be useful to set them out. The essential

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 6

Comparison of Composition and Recent Movement of Various Published Price

Board of Trade Index for basic materials (excl. fuel) for non-food manufacturing industry		*Financial Times* Index of sensitive world commodities		*Statist* Index	
Date	30th June, 1949=100 converted to July, 1952=100	Date	July 1st, 1952=100	Date	1867–77=100 converted to July, 1952=100
1952		1952		1952	
September	97·2	September	95·5	September	97·2
December	96·3	December	93·6	December	97·0
1953		1953		1953	
March	94·8	March	92·0	March	95·2
June	91·3	June	86·0	June	90·0
September	89·6	September	85·2	September	88·5
December	89·0	December	85·9	December	89·1
1954		1954		1954	
March	89·0	March	91·3	March	88·9
June	91·1	June	95·0	June	87·8
September	90·2	September	91·0	September	88·5
December	91·9	December	91·3	December	91·1
1955		1955		1955	
March	94·8	March	90·6	March	92·6
June	94·6	June	91·4	June	91·5
September	97·9	September	91·0	September	90·6
December	97·5	December	90·4	December	

Composition of Board of Trade Index (i)		Composition of *Financial Times* Index (ii)		Composition of *Statist* Index (iii)	
Commodity group	% weight	Commodity group	% weight	Commodity group	% weight
Food and feeding stuffs	*Nil*	Food and feeding stuffs	33·4	Food and feeding stuffs	33·8
Ferrous ores and metals	6·0			Ferrous metals	5·4
Non-ferrous ores and metals	13·0	Non-ferrous metals	25·0	Non-ferrous metals	5·4
Mineral oil products	5·0			Mineral oil products	5·4
Coal for carbonizing and benzole	6·5			Coal	3·6
Other mineral products (not elsewhere specified)	2·5				
Textile materials	31·5	Textile materials	25·0	Textile materials	19·6
Timber	14·0			Timber	3·6
Paper-making materials	6·0				
Oils, gums, waxes, etc. (other than mineral)	6·0	Edible oil	8·3	Oils, oilseeds, etc.	8·9
Hides and skins	4·5			Hides and skins, leather	8·9
Rubber	3·0	Rubber	8·3		
Misc.	2·0			Miscellaneous	5·4
	100·0		100·0		100·0
(i) Based on over 100 price series for 76 different commodities.		(ii) Based on 12 commodity price series.		(iii) Based on 56 price series for 45 different commodities.	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Indices for Raw Materials

Reuters Index		*Moody's* Index	
Date	September 18th, 1931=100 converted to July, 1952=100	Date	December 31st, 1931=100 converted to July, 1952=100
1952		1952	
September	97·6	September	97·6
December	94·1	December	92·5
1953		1953	
March	93·1	March	96·1
June	90·7	June	94·9
September	88·8	September	94·8
December	87·6	December	93·3
1954		1954	
March	89·2	March	98·2
June	88·9	June	98·9
September	88·3	September	93·5
December	90·8	December	93·4
1955		1955	
March	90·9	March	90·6
June	90·5	June	92·4
September	89·4	September	93·5
December	89·3	December	92·5

Composition of *Reuters* Index (iv)		Composition of *Moody's* Index (v)	
Commodity group	% weight	Commodity group	% weight
Food and feeding stuffs .	42	Food and feeding stuffs .	48
		Ferrous metals . .	10
Non-ferrous metals (incl. silver) . . .	18	Non-ferrous metals (incl. silver) . . .	10
Textile materials . .	25	Textile materials . .	24
Oils, oilseeds, etc. . .	10		
		Hides	3
Rubber	5	Rubber	5
	100		100
(iv) Based on 21 commodity price series.		(v) Based on 15 commodity price series.	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

difference between the currently published Board of Trade "basic materials" index and the other price indices referred to, lies in the exclusion of food commodities from the Board of Trade index. The Board of Trade index seeks to measure the price movement of "basic materials (excluding fuel) purchased by non-food manufacturing industries in the United Kingdom"—and the price series for the various materials used in compiling it are combined in proportion to the value of these materials entering non-food manufacturing industry in 1948. By contrast, the other indices are, with the exception of the *Statist* index, more concerned with illustrating the price trends of commodities entering world trade—and rely on a far smaller number of quotations.

(*b*) The *Moody* index is an American one, the weights being determined in accordance with the value of total U.S. production and/or consumption in the five years preceding 1932, and the prices used are based on American market quotations. It has much the same price base (December 31st, 1931) as the *Reuter* index (September 18th, 1931) and 9 out of the 15 commodities priced are also included in the latter index, although in this case the prices are mainly c.i.f. United Kingdom. The *Financial Times* index includes both dollar and sterling area commodities, the prices used being, as far as possible, those prevailing at the recognized centre for world trade in each commodity. The choice of items has been governed by what is judged to be their sensitivity and importance in total world trade. The *Statist* index is comprised of 45 commodities and is perhaps the best known of all these indices showing, as it does, the trend of major raw material price movements over a very long period. The *Moody*, *Statist* and Board of Trade indices all employ the arithmetic mean for averaging price relatives, whereas the *Reuter* and *Financial Times* indices both use the geometric mean.

(*c*) For the purpose for which these various indices are constructed, each is doubtless extremely useful, but it is certain that differences between them and the index for basic materials currently published in the *Board of Trade Journal* need to be kept in mind if confusion is to be avoided. For interest, details of each, converted to a common (July, 1952) price base, are given at Table 6. In view of the questions that have been addressed to the Board of Trade on this matter from time to time the information contained in this table may be useful.

(7) *The Reliability of the "Laspeyres" Formula*

50. The difference that has resulted from the use of the Laspeyres Formula in calculating the new series of index numbers, as compared with the use of (say) the Paasche Formula, cannot yet be adequately assessed because of the lack of suitable current weighting data. A roughly calculated set of 1954 quantity weights, for the basic materials index, however, produces the result shown at Table 7 below. This suggests that the upward bias caused by calculating on the Laspeyres basis is sufficient to emphasize the case for changing to a more up-to-date base as soon as possible, and, as mentioned earlier in this paper, a change to the 1954 = 100 base will be made as soon as the results of the 1954 Census of Production have become available. As has long been recognised, the bias in this kind of index does, if prices are rising and weights are changing significantly, tend to grow with time but it is always a difficult matter to weigh against each other the advantages of continuity on the one hand and more relevant weights on the other.

TABLE 7

Comparison of movement of Basic Materials Index Between June 30th ,1949, and December, 1954, using Laspeyres and Paasche Type Calculations

Calculation	June 30th, 1949	Average, 1954	December, 1954	Percentage Change June 30th, 1949, *to* December, 1954
Laspeyres (June 30th, 1949=100) —with 1948 quantity weights	100·0	143·5	146·4	+46·4
Paasche (1954=100) —with 1954 quantity weights	72·63	100·0	101·98	+40·4

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX I

The Questions Asked and the Problems Met in the Process of obtaining the Co-operation of Industry in the Collection of Prices

(a) *The Objects of the Index Numbers Explained*

(1) In approaching industry, it is essential to be clear as to the objects one has in mind in constructing index numbers. From the official and academic viewpoint, the main objects of the new series of wholesale price index numbers can be briefly described:

(i) To provide information about changes in the price structure.
(ii) To assist in deflating value series (though there are particular difficulties in constructing wholesale price series designed for this purpose).
(iii) To help in price forecasting and the understanding of shifts in the price structure.
(iv) To make possible the building up of a body of information which economists and social scientists in general can regard as raw material for longer term economic analysis, the fruits of which cannot be forecast but which, like the Census of Population, etc., can be regarded as an essential basis for a fuller understanding of the workings of the economic system as a whole.

(2) In addition there are the various commercial uses to which trade associations themselves can and do put price index numbers, for example:

(i) General market intelligence.
(ii) Appreciation of the trend of prices for materials entering various competing trades.
(iii) For comparing specific product price movements, either with average price movements or the price movements of similar products.
(iv) For comparison of particular and general price movements in the United Kingdom and in competing exporting countries.
(v) For use as "Official" price measures, to which reference can be made in wage negotiations when discussing the trend of "Costs" (e.g. the price of capital goods and of materials purchased).
(vi) For contract price adjustment purposes—as an alternative to maintaining detailed accounts on the price movements of every single item entering into a given long term contract.

(b) *Sampling: Selection of Representative Firms and Commodities*

(3) A sampling problem has to be faced in collecting prices for use in constructing wholesale price indices, since the range of producers and products that can be covered is severely limited by official resources and the response that can reasonably be expected from industry. The size of the sample of producers approached will, of course, vary considerably according to the type of product and the structure of the industry, and it is always valuable to obtain the advice of the Trade Associations as to which firms would be most likely to co-operate in providing prices or whose resources are best suited to this task. It is also essential to consider the extent to which the firms selected as contributors are likely to produce representative "price movements" common to the same product throughout the trade. In the great majority of trades it is fortunately found that the dispersion of the price movements for similar commodities quoted by different firms is relatively slight; the accuracy lost by accepting a comparatively small sample thus usually proves small, but every effort has to be made to watch for the exceptional trades where there is a real dispersion between the price movements of commodities produced by different firms and to try to increase the size of one's sample in such trades. For the furniture trade, where there are both quality and mass produced goods, for example, one must try to ensure that both are suitably represented. In industries where the membership of trade associations is not necessarily all-embracing one must endeavour to obtain satisfactory advice as to the main firms outside trade associations. Finally, one must consider the extent to which the data obtained must be given some special weighting treatment to allow for the absence of price quotations from a large number of producing firms and for the relative significance of those represented.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(4) The range of trade products which should be represented in the index is normally not difficult to agree, since one is able to present as a basis for discussion with trade associations a price collection framework based on Census of Production data and on the specialized knowledge of officials conversant with the trade. It is always advisable to concentrate on the commodities that one knows should be represented in the index, rather than to devote energy to the collection of prices for very minor items which it is known can safely be ignored, but this is primarily a problem of weighting as indeed is the whole question of commodity coverage of each index. The aspect of this matter which introduces the question of "sampling" lies not in the choice of commodities as such—for which an effort is made to achieve complete coverage, but in the fact that there is normally a very wide range of specifications of particular commodities which are sufficiently different for it to be necessary to discuss with trade associations the range of specifications that need to be priced in order to get a representative body of information together on the price movement of the commodity group as a whole. Here again practical problems become dominant, since one is necessarily limited to obtaining prices for a reasonably manageable number of items and one must necessarily take the advice of those in the trade who can be expected to know most about the matter.

(c) *Specifications of Commodities Required*

(5) It is normally desirable to determine, in the preliminary discussions with trade associations, (a) the "broad" specifications for which prices should be requested from individual firms and (b) usual terms of sale—e.g. what are the normal "lots" quoted for, if there is any question of "returnable bags or containers", and whether much business takes place on period contracts, or if there are normal discount terms which should be known? This sort of information helps to ensure that in writing out to individual firms one uses wording that is meaningful, leading to as little misunderstanding (and thus correspondence) as possible.

(6) It is not, of course, possible or desirable to suggest to each manufacturer the precise specification that should be quoted, but he can, however, be asked to state the fullest possible specification of the item (of a given kind) that he considers to be broadly representative of his output of *that class of commodity*, together with his normal terms of sale. The importance of asking that the item selected be fully specified, both in respect of physical attributes and normal terms of sale, is not to be under-rated. When, at a later date, one wishes to make sensible comparisons between commodities of a similar kind for purposes of substitution within the index (or even, for example, as between different countries) an insufficiently minute description (e.g. the width or breaking strength of rubber belting) may make this completely impossible or very misleading, and thus lead to discontinuity in a series.

(7) As regards "terms of sale" over a wide range of products there are comparatively small differences, in the United Kingdom, between the delivered and the ex-works prices and price movements of most commodities, the exceptions mostly being heavy and bulky materials (including a number of building materials). An excellent example is that of "granite chippings" which increased in price by 20 per cent. (ex-quarry) over the period June 1949 to June 1955, while increasing by more than 40 per cent. (delivered London) owing to the fact that this item comes mainly from remote (Westmorland and Scottish) quarries. For the great majority of commodities, however, the difference between delivered and ex-works prices is small, whereas the difference between the movements of delivered and ex-works prices is generally negligible because although e.g. rail freight rates have risen heavily since June, 1949 (up about 75 per cent.) they normally carry only a small weight in the delivered price. By contrast, in a vast country like the United States, the importance of freight is greater and there is a considerable difference between the movements of, and the absolute levels of, the prices of delivered and ex-works commodities. In the United States it might matter very much indeed which type of price one was dealing with if comparisons were being made. Despite what is said about small differences, however, it is conceptually desirable, in the United Kingdom as elsewhere, to use "delivered" prices for input price (materials) indices, and "ex-works" prices for most output price indices.

(8) The accuracy of price indices depends to a great extent upon the thoroughness of the physical specifications of commodities provided by those who contribute price information. Price indices become much simpler to construct and more reliable, for example, in proportion as "British Standard" descriptions are adopted by manufacturers—and this applies especially to chemicals and

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

textiles. It is, however, a fact that quite a number of British manufacturers do not provide more than a summary description of their commodities on their ordinary trade price lists and that when it comes to scientific descriptions, involving quantities or volumes, a number of commodities of British manufacture (though not engineering ones, which are usually very fully specified) are only superficially described. Despite this lack of adequate "public" specification of some commodities it is, however, usually easy to obtain precise guidance from manufacturers as to the actual content of the commodity one is interested in and to make suitable adjustments to the price series for that commodity when changes of specification occur. It is certainly essential to be kept fully informed of changes, however small, in the "specification" or "terms of sale" of any commodity and most manufacturers co-operate wholeheartedly in providing this kind of information, taking pains to be precise about technical changes which are taking place in the specification of any commodity the price of which is notified to officials regularly.

(d) *The Acceptability of the Weighting and Price Base of the Index*

(9) The price base for the new series of index numbers was taken as June 30th, 1949. The use of a single day as the price base was at one time thought likely to create some difficulty; in practice this arrangement has proved both convenient and satisfactory. Only in the case of a few materials was it found that the price at June 30th, 1949 was significantly different from the average for July as a whole, copper, zinc and brass ingots perhaps being the only important examples. Moreover, the use of this single-day base enabled us to avoid the biasing effect of including pre- and post-"devaluation" (September, 1949) prices in our price base period, as would have been the case had we used an average for, say, 1949 as a whole. The choice of this single day base date for prices greatly simplified our price collection problem and it is highly doubtful whether, if we had set our demands higher for price information from the manufacturing firms concerned, we should have obtained the co-operation we did.

(10) Our decision to adopt a weighting system based on the pattern of sales in 1948 proved acceptable to most trade associations, once it had been explained to them that this was the year in which the latest full Census of Production had been taken and was hence the only period for which data were reliable over the whole field we wished to cover. It must be appreciated, however, that in the case of some manufactured products our approach to trade associations could not be made until 1951 or 1952 and trades in which there had been a significant rise in sales, or radical changes in the pattern of production, since 1948, were naturally in favour of a more up-to-date pattern of weights being used. While in general the change in the pattern of production and sales was not significant between 1948 and 1951, there were some cases where it was clear that the 1948 weights were unrepresentative if not grossly misleading. This was particularly true of commodities where rationing ended after 1948, as in the case of soap, or where new techniques were being rapidly developed, as in the case of nylon and rayon products, or where production was increasing very rapidly, as in the case of television sets. In nearly all these cases it was possible to be flexible so far as a few commodity indices were concerned, or even so far as narrow trade sector indices were concerned, without in any way departing from the basic principle that the weight carried by the particular commodity or trade sector index in a major composite would nevertheless be a 1948 one. Thus, while it was deemed reasonable, at the request of the trade, to give soapless detergents a 1951 weight of 20 per cent. in the narrow trade sector index for "soap, candles and glycerine" (the 1948 weights being much smaller) the value weights given to soap and soapless detergents in the index for "all manufactured products" were 1948 ones. Publication of this index was considered desirable since it covers a Census of Production "trade" grouping—one which corresponds with S.I.C. minimum list heading 35/1.

(11) There are so many approximations made because of sampling or for other reasons, in calculating any index, that to be purist about maintaining 1948 base weights in indices for the few trades where change has been rapid, would be quite unrealistic and might result in a refusal from firms in the trade to co-operate in providing prices for the purpose of constructing an index that would be of little meaning to them. In most cases, moreover, even though the effect on a particular commodity index of using up-to-date rather than 1948 weights might be significant, the effect on a composite index in which it featured would be relatively insignificant. Accordingly, provided that the exceptions were few and far between, the making of them was desirable rather than the contrary.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In the majority of such cases, moreover, it was possible to come to solutions which made it possible to avoid the problem presented by 1948 weighting. For example it was not essential to publish a composite which combined radio sets, radiograms and television sets, using 1948 weights, since one could publish (as we did) the three indices, separately, as commodity indices. (The fact that 1948 weights were used in combining these in some larger composite did not affect the interests of the trade).

(e) *Standard Products*

(12) The problem both of weighting and of base pricing is, of course, linked with the problem of finding suitable "standard" products to price. In the few trades where there is no such thing as a standard product (or anything approaching it) since the major products of the industry change very rapidly, as in the aircraft industry, the problem of constructing a price index is very difficult. The aircraft industry is the only major one where the construction of an index has so far proved quite impracticable, and where it would indeed be very difficult to calculate even a chain-based index, since, because of the method of pricing aircraft, it is difficult to find a basis where comparable prices can be obtained even over short periods. The same sort of problem is, however, met, to a lesser degree, in preparing indices for other products (e.g. buildings) and it is a potent cause of approximation. Where changes in design technique and scientific inventions are so numerous and so continuous that representative commodities are hardly even comparable from one quarter to another, index numbers cease to be calculable. Fortunately, this is only true of fairly narrow sectors of industry since, normally speaking, these sort of changes take quite a period to make themselves felt and are introduced in easy stages. It does, however, mean, that in any trade where tailor-made commodities are produced in greater quantities than "standard" items, or in significant quantities, as is true of the machinery, castings, forgings, etc., trades, the constructor of any price index has to assume that these missing prices move in close conformity with those of some more standard (but similar) products. Such approximations may sometimes be significantly incorrect.

(f) *Confidentiality of Prices*

(13) In approaching trade associations it is essential to offer to treat as confidential any prices that they or their members might agree to provide. This applies more particularly to the prices of finished products, as prices of most raw materials and many semi-manufactures are widely published or known. As prices of manufactured products are, in the main, collected from a sample of firms only, there is a danger that prices quoted as providing a reasonably representative guide to price movements might come to be quoted as "representative average prices" for that particular commodity. The use of such prices without consulting the firm or trade can lead to serious misunderstandings affecting the interests of the trade and creating extra work for those who have to give explanations and justifications. This is because a sample of prices that is representative of "price movements" may not be at all representative of absolute prices. For many commodities, individual prices can subsequently be published or revealed by agreement with the individual firms, but the circumstances in which such prices are revealed are very important and it is often desirable that the form in which they appear should be agreed with the firms or trades concerned. In this connection it will be noted that the Board of Trade has in recent years been publishing annual prices for a selected range of commodities in the annual review of wholesale prices which normally appears in February.

(g) *Simple Procedure of Obtaining Regular Price Information from Firms and Trade Associations*

(14) It is essential to have a simple but flexible system of price collection which can readily be explained to contributors. For the great majority of trade associations and firms a specially designed shuttle reply-paid price quotation card, allowing for the insertion of code numbers, specifications, terms of sale, and details of price changes *with their dates*, proves readily acceptable. Such cards, even if they relate to quite a number of commodities, can normally be handled easily by the clerical staff in the factory or office concerned and can be sent back to the Government Department concerned in the last days of each month.

(15) Experience has shown that this simplification of the work of reporting price changes is welcomed by firms and is helpful in ensuring both continued co-operation and care in reporting.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The bulk of our price information, indeed, particularly for finished goods where prices rarely change more than once a month, is collected in this way. In suitable cases the shuttle price quotation cards may be sent out less often than once a month, or contributors may elect to provide prices, in the form of a letter or statement, at monthly or quarterly intervals, or, where price changes are exceptionally infrequent, only when a price change actually occurs. Where prices are liable to fluctuate daily throughout the month, as do those of many non-ferrous metals, it is sometimes possible to obtain an average monthly price from the contributor rather than details of each daily price, since information of this kind is often maintained by the trades concerned. It should be stressed, however, that where price information is obtained only at intervals it is desirable to obtain details of the date from which any new price became operative.

(16) In a few instances price information is received already expressed in index number form. This happens, for example, where firms prefer to pass information through their trade association. The construction of these indices, so far as internal weighting and composition are concerned, is agreed with the trade associations.

(h) *Subsequent Co-operation with Trades*

(17) It is always necessary, from the outset of one's conversations with trade associations, to make it clear that price indices calculated for sectors other than their own will, so far as possible, also be made available to them. It is possible to promise that when indices are calculated and published they will readily be made available to each trade and each firm assisting, in the form of reprints of articles appearing monthly in the *Board of Trade Journal*. These reprints are, in fact, sent out free of charge to all our contributors, who can in this way keep in close touch not only with those indices to which they contribute, but also with the other published indices. Within the limited resources available we also try to meet any requests from contributors (and others) for unpublished (but non-confidential) information. Furthermore, any request from a trade association for calculation of a price index likely to be particularly useful to them is carefully considered. Thus, the help they provide in one direction may well be returnable to them in another.

APPENDIX II

The Storing and Processing of Price Information and the Calculation of Price Relatives and Index Numbers

(i) The method of obtaining price information and the necessity for a simple procedure has been dealt with above. Having obtained the prices, too much stress cannot be laid on the importance of efficient methods of recording this information and processing it.

(ii) In the Board of Trade system each item has a code number based on Standard Industrial Classification Minimum list headings. The date of receipt of each price quotation card is recorded and the cards are passed to a calculating section where, after examination for changes in specification, etc., the prices are copied on to working sheets against the appropriate code number, for conversion to price relatives on the basis of June 30th, 1949 = 100. Each copying and calculating operation is done independently by two clerks, one calculating an average monthly price and converting this to a price relative, the other converting each daily price to a relative and then calculating an average monthly relative. These answers are called over before each begins the next step of combining price relatives into index numbers. It has been found that unless some such checking method is adopted, the system is in danger of coming to grief because of errors which could not otherwise be easily detected or traced, especially as most of this calculation work has to be compressed into two or three days after the end of each month.

(iii) Some of the resultant index numbers will necessarily have to be shown as "provisional" on first publication, but the period following the calculations is devoted to (*a*) examination of the individual cards received to see that the information provided on them is consistent; (*b*) writing to firms who have failed to send in their cards; (*c*) carrying out such correspondence as is necessary either to seek the correct information on an item where the information provided appears (on the face of it) incorrect, or to answer queries raised by contributors about quality changes or the going

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

out of production of particular items; (*d*) amending records to agree with all changes of specifications notified; and (*e*) recording on permanent record cards the average prices and relatives calculated for the latest month.

(iv) It is essential that price information is held on permanent records which can be quickly and easily referred to. Our method is to keep two record cards for each item, one recording prices, another the corresponding price relatives. These cards are stored in "visible index" folders leather bound to withstand constant handling, so arranged that all cards relating to one particular index appear, as far as possible, in the same folder. Specification and code numbers and the weight the item carries in its immediate index, appear on each card. The card relating to an index is distinguished by a coloured border from those for individual items. Small coloured slips attached to item cards indicate any composite index in which they might appear.

(v) Prices relating to several thousand commodities are received and converted within a few days through a whole system of working sheets into price relatives and index numbers, and then subsequently maintained so that they are readily available for all those who require them. The success or failure of such a system of index numbers (from the point of view of accuracy and of publishing on a regular and early date each month) depends to a very great extent on the proper organization of the price receiving and price calculating sections. Our index numbers sometimes appear in print within 11 days of the end of month to which they relate, almost always within 18 days, and this means that they sometimes go to press only a few days after the prices on which they are based have been received. Because of the great importance of speedy publication, as well as for other reasons, there is little doubt that it is wise, from the commencement of any price index system, to bear in mind the ultimate possibility of electronic processing of the data. To this end a system of code numbering the families of prices according to the indices into which they are fed should be organized in close consultation with machine computing experts. The greater the range of prices covered the more economical, practicable and efficient a process of electronic calculation will tend to be.

Discussion on Dr. Phillips's Paper

Professor C. F. Carter (in proposing the vote of thanks): It is a great pleasure to congratulate Dr. Phillips on his paper. It is five years since that distinguished leader of the band, Mr. Stafford, indicated in a learned paper to the Society that he was about to take the stage as Jack Stafford and his Wholesale Serenaders; and it is good now to be able to hear from one of his principal instrumentalists of their successful progress.

The old Cost of Living index is born again as an index of retail prices; but the old wholesale price index is dead, and its place has been taken by a whole range of sector input and output indices. I regret that, in this welcome change, the Board of Trade did not see fit to scrap the word "wholesale"; firstly because I would prefer to keep that word to attach to the prices at which retailers buy things (which is, I think, its common sense); and secondly because the change of title would have drawn attention to the complete difference in the concept and intention of the new indices. Incidentally, an index of the prices at which retailers buy things would be useful; if I understand Table 5 right, the "wholesale price index for manufactured personal consumer goods" represents for the most part two or more distributive stages before the retail price index, and indeed (oddly enough) is a measure of the prices at which wholesalers, if they exist, *buy*.

I hope that in due course we shall get a Blue Book to tell us about the new indices in precise detail. The two papers still leave a number of important gaps; for instance, we do not know which series are "unpublished but available to the public on request", and consequently we do not know what we can ask for. Ideally, one would like a complete list of the sector and commodity group indices, with their weights (so that one can combine them together to one's own fancy), and with an indication, where appropriate, of the reasons for confidentiality—i.e. whether the price series is secret, or merely suppressed to avoid the disclosure of some other series. I must congratulate the Board of Trade on only having to keep quiet about 40 per cent. of its manufactured goods prices—it would be nice to know what this means in terms of weights in the "all manufactured products" output index—but the bad habits of British industry in demanding unnecessary confidentiality need frequent exposure.

Of the many interesting technical points in the paper I will select one—the problem of seasonality. If a commodity, for natural reasons, is wholly unavailable for a period, the only logical price to set is an infinite one; and the only way to avoid having an index which runs away to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

infinity is to take account of the substitution of other commodities or the deferment of purchases—in other words, to use variable weights. But if we look at the matter further it can be seen that the difficulty of buying new potatoes in December or Christmas cards in January is not really distinct from the difficulty of buying pig-iron on Sundays. We take it for granted that an average price for a month is an average of selling days, whatever they happen to be; similarly there is no difficulty in defining a yearly average price for Christmas cards as an average for all the days on which significant quantities of Christmas cards have changed hands. I therefore suggest that, in all those sector indices in which seasonal commodities are important, the calculation of a yearly index should be conceived as something more than the averaging of the monthly indices. The various "tricks" used in dealing with seasonal commodities in a monthly index are not needed in a comparison of annual average prices, and this is therefore a case in which one could have an annual index different from, and better than, the monthly averages.

In general I find myself in hearty agreement with the techniques chosen for the new index, and for the ways in which the rigidity of theory has been modified for practical reasons—e.g. in departing from 1948 weights on occasion, though surely we should have a list of these departures. But the new indices are not yet complete; we still need series for the input and output of home agriculture, for the "consumption" of building, and other interesting prices. I am not quite sure how far the powerful soloists of the Ministries of Agriculture and Works and Fuel and Power are willing to play in harmony with the Board of Trade ensemble; but I would repeat a point made in the discussion on Mr. Stafford's paper, that it would be well to conceive the whole system of price indices, not excluding the indices of retail prices, as a single system under a single technical direction, and published in one place.

Mr. H. S. Booker (in seconding the vote of thanks): Dr. Phillips should be congratulated on an adequate picture of what the Board of Trade is attempting in its indices of wholesale prices and how it is doing it. It is satisfactory, too, that five years after Mr. Stafford told us of the task, Dr. Phillips can report, in effect, that it is completed. The next phase is repeating the job with a new series of weights appropriate to 1954. Part of the clearing up operations will be the post-mortem on the index from 1949 (or 1948 for weights) to 1954 of which chief evidence will be the calculations of indices on weights appropriate to each year. It is interesting to note that the basic materials index shows a rise of 46 per cent. as calculated on weights appropriate to 1948 but this would be only 40 per cent. with weights appropriate to 1954. The difference is stated to indicate an upward bias in the index as calculated, but this is unfair to the index; the difference in the indices is a measure of the adaptability of British industry to price stimulus.

The problem of the base period for prices was urgent in 1949 and a date (June 30th) was taken to avoid complications as a result of devaluation later that year, closely followed by the Korean boom. Devaluation should have a one-way influence but the Korean boom should be associated with a rise and subsequent fall. The two influences together may make it inappropriate to generalize much on the basis of Dr. Phillips's measures of price volatility. The base is very important:

	June 1949	*Average* 1950	*Average* 1955
As used:			
Basic materials . .	100	138·6	152·3
Manufactured products .	100	106·8	129·5
With 1950 *as base:*			
Basic materials		100	110
Manufactured products		100	121

As used, the indices show British industry doing well in relation to increased prices of materials. With 1950 as base they show the benefits of devaluation gradually being lost.

The Board of Trade has persuaded us that a single index number cannot be reasonably weighted if it contains both input and output prices, and, quoting Keynes, does not attempt a single "all articles" index. The problems of the Ministry of Labour are different; they are only concerned with finished goods consumed in the United Kingdom but they remain loyal to a single index comparable, perhaps, to the index of manufactured products and hence not giving rise to problems of "net weighting"—but nevertheless the emphasis is different. The Board of Trade is interested in numerous sector indices which can adjust money values to obtain quantity inputs and outputs or alternatively adjust quantity figures to obtain money values. The Ministry of Labour seems to make a mistake in not producing sector indices of retail prices appropriate to different income levels, family responsibilities, and even consumption habits (e.g. contrasting consumers of tobacco

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and alcohol with others). I am not, however, convinced that almost every other commodity can have an index of price except the commodity money. It seems a narrow view to avoid stressing the general level of prices as a unit of reference as if everything had been said when each sector's values had been deflated to obtain something called "real". There are two things we are interested in, the change in some "average" level of prices and the individual variations about or relative to that average. Changes in relative prices are as real as changes in volume—many people would feel there are advantages of working in an industry where output prices are rising relatively rapidly; between the wars certain industries not much affected by foreign competition were able to maintain prices better than others, people reaped real benefits from working in such "sheltered" industries, benefits which should not be deflated away by reference to a price index only appropriate to that industry.

To my mind much of Part II of Dr. Phillips's paper, which is concerned with this problem of relative prices, suffers because there is no accepted "average" as a standard of reference. Dr. Phillips has had to invent a standard of reference, but he has kept well away from any kind of "middle" or average value for this purpose; instead he uses the extreme and calls that 100. In Table 3 the price index for raw wool at the peak is called 100, the price index of wool clothing at the peak (with June 1949 = 100) was 30 per cent. of the index for raw wool and is so recorded as if, in some sense, the price of wool clothing was only 30 per cent. of the price of raw wool.

I have a preference for relating sector and commodity indices to some average, even if the average is not quite "pure". But even here the price level of output of industry is pure (that is correctly weighted) and might do satisfactorily if food and domestic fuel were added. Thus I am in favour of a wholesale price index which is generally accepted as being basic, if only because it indicates in a period of inflation that something can be introduced to the economy generally from the side of money. That is where the single composite index of the Ministry of Labour scores over the complex series of the Board of Trade.

On a matter of detail the sector index numbers we now have are numerous, but there seems to be a gap in transport and some other non-manufacturing industries which nevertheless use materials in providing services where indices, at least on the input side, would be useful.

The vote of thanks was put to the meeting and carried unanimously.

Mr. STAFFORD: I should, I think, make some kind of public apology to Dr. Phillips for what has been a very unequal sharing of the relative burden. It is tolerably easy to say what one thinks should be done; it is most difficult to go away and do it. Dr. Phillips is much to be congratulated both for this paper and for the work which underlies the experience he has given us tonight. He has worked very hard over the course of years to prepare the information which is given in this paper and which appears, printers permitting, month by month in the *Board of Trade Journal*.

I should like to add two small things by way of postscript to what Dr. Phillips has said. The first hangs on his plea for the Civil Service Commissioners to recruit some of the extraordinary mortals who can cope with these index number problems. There is, I think, a moral to be drawn from paragraph 32. When compiling index numbers of wholesale prices this problem of getting a continuous measure over time is one that you cannot escape. People write and say they have stopped making certain models and ask what they should do now. That is the sort of thing to which a solution has to be found. Therefore, what may in concept be a relatively simple job becomes extraordinarily difficult because of the discontinuities.

This is not peculiar to wholesale price index construction. It is something which every economic statistician faces and with the compilation of almost every index number it peeps its head out. The same difficulty is met with in compiling index numbers of export and import prices, but there the problems tend to be masked by the continuity of the Export and Import Lists—rather different goods can be placed in different years under the same heading. There is a moral that these problems pervade index number construction. They underline the fact that all these things are quite brittle and should be used with great care.

Experience in using our system of wholesale prices and conversation with people has underlined the difficulty of producing a system which is both summary and helpful. It is very difficult to strike a happy balance between summarization which is misleading and a mass of information which is indigestible. I would part company with Mr. Booker on this and would favour publishing as much detail as we can, giving as much information as we can, producing as many useful weighting systems as we can, and letting the user compile that which is apt for the purpose.

Miss M. DEANE: I should like to add my congratulations to Dr. Phillips for his very lucid and interesting paper. To me the most telling reflection of the great improvement in the Board of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Trade Index of Wholesale Prices is that it is only a matter of four years ago that *The Economist* dared to compete with the Board of Trade and produce its own index. We have now discreetly buried this and I am sure no private investigator today would conceive of doing the task or would see any point in attempting it.

I agree with Mr. Stafford that the more that can be published about the methods of constructing these index numbers the better. I do not think it is a good thing to try to hide the peculiarities and I think the time has come for bringing the limitations of index numbers into the open and emphasising to the layman the sad truth that in these calculations there is no single perfect formula. The different weighting systems can produce a difference that cannot be ignored. I think it is a good idea to publish the full particulars and then to let investigators, if they wish, try to make some cross between the different formulae.

The statement which grieves me in this paper is the one in paragraph 4, that only 60 per cent. of the output indices for manufactured goods are published or can be obtained from the Board of Trade. The missing parts I think cover almost the whole of engineering, ship-building and vehicles. I believe I am right in saying there are no output indices for any kind of machinery or machine tools. It is not the fault of the Board of Trade but due to the reluctance of the industries concerned to let the information be disclosed. I think any help this Society can give to stimulate industry in showing the usefulness of publishing these index numbers should be given.

About grouping price relativities, would it be possible to divide between imported and home produced materials? One of the uses to which economists will undoubtedly put these much more refined wholesale numbers is in trying to observe inflationary tendencies at an early stage. Equally useful on the output side I would say would be a division of manufactured consumer goods between durable and non-durables.

Another point concerns measures of volatility. In paragraph 46 we are told that the only two major commodities whose price movements, in the five years ending December, 1954, have changed direction four or more times, each time by more than 20 per cent., are raw rubber and raw wool. I presume the measurement of 20 per cent. is "from peak to trough"? I rather feel emphasis on this characteristic might give the wrong impression of the predominant course of these price movements since the changes in direction lasted in most cases for only a month. I am not sure that I see much to be said for distinction between prices which remain absolutely unchanged and those which move by a fraction of a percentage. What I think is lost with these measures is the fact that rubber took considerably longer to fall in price, a characteristic which I do not think is brought out.

I am sure that any such analysis as this of price volatility that the Board of Trade can make will be extremely welcome.

Professor ALLEN: I do want to welcome the publication of this paper. Dr. Phillips says he hopes the paper will be useful to those in this country who use the indices and also to those in other countries who are thinking along our lines and might learn from our mistakes and experience. I should like to underline this and, indeed, to extend it. There are people in other countries who are not thinking on our lines and who might still be converted to the proper way of thinking. I know perfectly well, having tried to "sell" the idea of these sector indices at United Nations seminars in other countries, that there are countries which still want a general wholesale number with a few broad constituents. These are not only less developed countries, unable to afford to put together hundreds of different commodity series, but also some of the more developed countries.

I find myself very much in agreement with the Board of Trade and opposed to Mr. Booker on this matter of the general or all-article index. I would not change anything I said in commenting on Mr. Stafford's earlier paper except perhaps to take out the qualification, made at the time, that a general index is all very well in one column on a page of United Nations statistics.

There is only one point of disagreement I have on Dr. Phillips's paper, and this may be more a matter of emphasis than of substance. I refer to the section, paragraphs 28 to 34, on the problem of quality changes. These paragraphs seem to be rather vague and I found it rather difficult to ascertain what was going on. My general impression is that Dr. Phillips, and presumably the Board of Trade, think that cost considerations are the true criterion, or at least the most practical criterion, for a link when there is a quality change or change in specification. I would disagree with this.

Take one example. A manufacturer of brushes switches over from ordinary bristles to nylon bristles and finds the cost of the brush with nylon bristles is, say, half; but the new brush is just as durable as the old, satisfies the consumer just as well, and the price is the same. If you link by cost you get the price doubled; if you link by durability the price is unchanged. As another example, take the case of a manufacturer who supplies containers. An improvement is made in

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the materials of the containers; the cost is the same but the product is twice as durable and is sold at the same price as before. If you link by cost the price is unchanged; if you take durability there is a fall.

I do not see why the cost criterion should be the one for the link. The index numbers relate to transaction prices and there are two sides to a transaction, the supplier's and the purchaser's. There seems to be no reason to take one side of the transaction rather than the other in linking; the criterion of cost is not the only one in making the link. This is the only objection I have to an otherwise excellent paper.

Dr. BARNA: I too feel unhappy about the paragraphs relating to quality. Two points are made to which I think exception can be taken. One is that the problem of correcting for quality changes in wholesale prices is different from the problem relating to retail price indices. I think the theoretical and practical problems are identical.

There is, secondly, the statement that if one looks at quality the question of subjective valuation comes in. I thould like to add to what has already been said that in a large number of cases subjective valuation need not come in. I refer to capital goods. With some effort, certain objective criteria could be found to measure the "quality" of capital goods. The problem of capital goods rather worries me because in most of the examples given by Dr. Phillips we have instances of manufacturers changing the raw material content of the commodity, and then a few weeks later perhaps changing back. In the sphere of capital goods, however, it is generally thought that quality is improving all the time and, if no account, or insufficient account, is taken of this, then the index is bound to be affected. In a typical case a new machine costs 10 per cent. more than the old but may be 50 per cent. more "valuable". A new machine may cost a little more but it may run three times as fast. I am sure some method could be developed in regard to that, although it would not be an easy one. The new machine would not necessarily be three times as valuable.

In this connection it is regrettable that the index numbers for engineering are not published, and it is impossible to judge how far the Board of Trade has taken these problems into account. The suppression of the information is not quite complete as one can work out the annual indices relating to plant and machinery from tables in the Blue Book on national income, but of course it would be preferable to have these indices published explicitly.

I have recently come across the problem mentioned above in the course of my work. I have met a number of manufacturers who were very much interested in obtaining a realistic value of plant and equipment and for this purpose they needed a good index number of prices. In most cases they either used *The Economist* index for capital goods, which cannot possibly contain as many items as the Board of Trade index, or they constructed index numbers of their own. I think it would be easy to demonstrate to manufacturers that they would greatly benefit if the Board of Trade were to publish these index numbers.

Mr. G. R. WHITE: Representing a trade association which provides Dr. Phillips with one or two ingredients for the pie, it has been a very great pleasure to me not only to see the pie when it has come out of the oven but also to see the recipe.

I was very pleased to see that the basis of the index was going to be changed from June, 1949, to 1954 and, in particular, to an average for that year. Personally I always mistrust indices based on one date, but I realise the difficulty with which the Board of Trade was faced when the post-war index was first started. 1949 was a year when many industries had not quite come out of control and when the price structure was still somewhat artificial. The 1954 basis will give a much more realistic comparison than the old basis, but it would be interesting to know whether there is any estimate of the proportion of the field covered by these indices which really was in an artificial state owing to control in 1949.

Mr. LEAK: In regard to Mr. Booker's table, it seems to me that we have a partial explanation in Table 3 of the paper, which shows that there is a very considerable time lag between the price of raw materials going up and the change in the price of the manufactured articles derived from them. If the second line of the table were time-lagged by six months we might see more readily what had happened.

The question of quality changes which has been referred to arises very acutely in the case of potatoes. These were in the old index and I imagine they will be in the new index in due course, but old and new potatoes are very different. Potatoes are potatoes, but one variety is of a better quality than another and commands a higher price. How during the course of the year is it proposed to cope with the problem of the changing quality of potatoes?

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Dr. J. WISE: There is one point that puzzles me in the section on quality changes. The author says, "For example, the use of an additional proportion of synthetic fibres in a wool cloth or the use of slightly different dyes are things which are difficult to assess, especially in the clothing field where the manufacturer or seller of the "garment" is probably not the manufacturer of the actual material from which it is made. There is, therefore, a danger that, sometimes, insufficient allowance for quality changes may be made and the movement of the index may, accordingly, be damped." I was not altogether clear why the movement of the index should necessarily be damped. As far as I can see it might well be exaggerated if the quality changes arose from changes in technique, tastes or income levels.

The author refers to cases such as variations in the fibre content of wool cloth. When we want to use price index numbers in some kind of demand analysis, to correct for such changes it would be a good idea to construct some kind of quality indicator, such as the percentage by weight of synthetic fibres in the cloth, and if such a quality indicator were used a great improvement might result in the estimates. In considering beer consumption Stone did in fact explicitly bring the specific gravity of the beer into the demand equation. Where quality variations are considerable, a quality indicator should be used if possible to relate the quantity and price series to approximately comparable units.

Dr. PHILLIPS subsequently replied in writing as follows:—

In reply to Professor Carter, the only output price indices "available to the public" that are not 1948 base-weighted are those for "soap, candles and glycerine" (including the sub-index for soap) and for "rayon fabric". Both are mentioned in Appendix I. On the materials side, the ending of controls and other such changes have led to post-1948 weights being used in the reconstructed series for certain materials including imported softwood (from April 1952) and hides and skins (from January 1955.) The unpublished electricity component of the Ministry of Fuel and Power published index for fuel used in manufacturing industry is, as stated in the *Board of Trade Journal* on 14th January, 1955, based on average values. All our unpublished composite indices are 1948 base-weighted, so far as their main component series are concerned, but a few of their sub-indices, including the three for small, medium and large passenger-cars, have internal structures that reflect a more recent pattern of sales.

We have already published the main group weights used in compiling the more important of our indices including that for all manufactured products. To give details of all the weights used in combining all our indices (regardless of the demand for this information) would involve undue expenditure of time. The difficulty in issuing these weights is not a clerical one, but arises from the fact that in every single case we have to make certain that the weights issued to the public are so set out as to avoid not only (a) the disclosure of confidential Census of Production data but also, (b) giving away the weights of confidential series, the movement of which could in consequence be calculated. In view of the staff available, we feel that we can only justify issuing lists of weights in answer to *ad hoc* requests. When such requests are received the weights for the index concerned are subsequently duplicated: there are at least 70 such lists which can be sent to anyone interested. There is little we can say about the indices that we cannot publish at all without breaking faith with whose manufacturers and trade associations who go to so much trouble to supply us with the necessary price data.

In general it can be said that those indices that are "available to the public" but not published in the *Board of Trade Journal* monthly, are either commodity group sub-indices of published composite indices, or composites of published commodity indices; they are not published mainly because we cannot find the space to print them, but also, in some cases, because we do not wish to suggest that their value is as great as might be implied by publishing them. A number (75) of these commodity indices do nevertheless appear in our annual article in the form of annual averages. Others are regularly provided in answer to *ad hoc* requests. A typical example is the index for hematite pig iron, one of a large number of commodity indices featuring in our Iron and Steel index which, though not published by us as indices, are obviously available to the public since they can be calculated by anyone from the price determination lists issued by the Iron and Steel Board. Bowing to Professor Carter's very understandable desire to know about the indices which though not published are available to the public, we are considering issuing a list of them.

We have considered the question, also raised by Professor Carter, whether it might not be better to drop the title "Wholesale" from our series and to use some other word. The Germans, as you know, call some of their indices "producers' price indices". While this is quite a good title for price indices of manufactured products it is quite unsuitable for materials, and hence also as a portmanteau heading. Another possible solution is the American one of calling these indices "wholesale *primary* market price indices", though the fact that the word "wholesale" is

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

used in qualification of "primary" makes this a doubtful advance. The Board of Trade will certainly give this matter further consideration before rebasing on 1954 = 100, and would welcome suggestions.

In reply to Professor Allen I would emphasize that we do not have any "rule" in the Board of Trade suggesting that allowances for quality changes should be made on the basis of the cost of such changes. On the contrary, as stated at paragraphs 30 and 33 of my paper, we think that each case must be "examined on its own merits" in consultation with the manufacturer/price contributor. The brush example raised by Professor Allen really amounts to the substitution of one commodity for another; we would only be justified in accepting prices for the substitute commodity—in our base-weighted index for brushes and brooms—on the assumption that bristle brushes were virtually displaced by nylon ones. Given such a case we would certainly not ask for any cost data; the change concerned is obviously far too radical for treatment in that way, being, as it were, a qualitative change in quality not a quantitative one. While, therefore, we *would* seek cost data in designing our quality link—if a rubber handle were removed from a pram (the previous price of which had included a rubber handle) or if a windscreen wiper were added to a car (the previous price of which has been quoted without such a wiper) the matter would have to be treated as one of doubt where the change was not so simple or so obviously quantitative. As stated in paragraph 32 of the paper we have tended, when in doubt, to make our link at the level of the last relative for the old item. In the absence of clear technical advice from our price contributor to the effect that some other treatment is fairer, there are few alternatives. This previous relative link would be our normal practice in such cases of substitution, but we would first make sure that other prices for the old items were not available and that the substitute item was regarded within the trade as a reasonable equivalent. I do not wish, however, to suggest that we always do the right thing, or even that we can always obtain enough information to be able to do the right thing, and can assure Professor Allen that we are always learning.

The sort of quality adjustments suggested by Dr. Barna are often impracticable, but he may rest assured that much thought is given to the quality adjustments we make to our indices; we do try to take account of changes in measurable quality characteristics—including the performance of machines—and always seek factual assistance from our price contributors to this end. Where quality is always changing the product becomes so non-standard that it may be difficult to obtain any willing price contributors: in such cases imputation of the weights for the items concerned to some more "standard" products of the same kind has to be considered, as stated in my thirteenth paragraph. When faced (as in motor cars) with dozens of quite important quality variables ranging from horse power and seating capacity to finish and comfort, each having a different significance in different models, some imperfect compromise with complexity has to be found. Fortunately, many products are comparatively easy to deal with and one such is beer, the specific gravity of which, mentioned by Dr. Wise, is, of course, fully allowed for in calculating our index for that beverage.

I agree wholeheartedly with Miss Deane's remarks about the difficulty of assessing "price volatility"; it was for this very reason that I decided on measure "A" as the best one on which to base my Table 4, rather than the change of direction measure. There are, however, many different aspects of price volatility and it would be difficult to dogmatize about any of them. I think that I have to some extent answered the point, raised by Miss Deane, about comparing the volatility of individual commodities with that of basic materials in general by my analysis in Section 3 of Part II of the paper, but I am sure that she is right in suggesting that my study, which was only intended to be illustrative, could profitably be expanded. Her point about distinguishing between durable and non-durable goods is one about which we would all agree. As may be seen from Table A, the Board of Trade maintains (and is planning to publish) a separate sub-index for household durables among its "personal consumption" price indices. The idea that we should distinguish between imported and home produced basic materials is attractive and we will certainly give this consideration.

Professor Carter has rightly pointed out that it would be possible to have an index that was both useful and meaningful on the basis of annual averages for seasonal agricultural commodities, even though no index that was useful from month to month were feasible. It would be better, he suggests, to face up to the desirability of having annual indices in this sector rather than monthly ones. We may be persuaded of the correctness of this suggestion.

Mr. Booker said that the relative movements of our indices for basic materials and manufactured products would have been vastly different had we used 1950 = 100 instead of June 30th 1949 = 100, but it has been one of the main objects of my paper to show how wrong it would be to choose a year like 1950 as a common base for two such indices—since both devaluation and the outbreak of the Korean War (June 1950) were reflected *much more rapidly* in the prices of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

materials than in the prices of manufactured products. In fact, the index for all manufactured products reached its peak in March, 1952, twelve months after the basic materials index reached its highest point. Our choice of 1954 = 100 for the rebased series has been carefully considered in the light of the fact that both these indices reached their post Korean low point in the same month, February, 1954, and moved comparatively little over that year. Mr. Booker also suggests that there is a case for comparing the price movements shown in Table 3 with the general movement of some overall indicator. It should be emphasized in this connection that it is not "prices" we are comparing, but price movements, and while I do not wish to suggest that there may not be some purpose in making such a comparison, I can see little value in comparing the price movements of finished products or basic materials against a reference line that is some sort of weighted average of both; an average, moreover, in which the determination of the weights would be most controversial. I most certainly believe, however, and develop this line of thought in my paper, that it is useful to make separate comparisons of the movement of individual basic materials and of individual products with the movements of the composite indices for basic materials and manufactured products, respectively. Incidentally, these comparisons between the movement of prices for individual materials (and products) and the average for basic materials (and manufactured products) have been a feature of annual articles in the *Board of Trade Journal* in recent years, (e.g. February 19th, 1955 and April 21st, 1956—Tables A–B). Even more useful than this, in my opinion, is the analytical approach adopted in Table 3, because this emphasizes the great *differences between trades* as to the relative movements of their basic materials and their finished products. By this kind of analysis one can see the extent to which the greater price volatility of certain raw materials—and the more significant role played by wage costs and productivity in one industry compared with another—tends to dampen the movement of prices of finished products (compared with the movement of prices of raw materials) more in some trades than others.

Dr. Wise referred to the possibility that quality changes could be made (as in the dye content of some materials) which might, if ignored, so far from "damping" the price index concerned, do the contrary. I used the wrong word and he is perfectly correct. I was in fact referring to a special case which perhaps is not uncommon in times of rising prices.

I should like very much to thank those who have spoken so kindly about this paper. It has owed much to Mr. Stafford who, as everyone here knows, is very much the father of these new price indices, and whose interest in the development of this paper has been an active one.

As a result of the ballot taken during the meeting, the candidates named below were elected Fellows of the Society:

John Dennis Biggers.
Richard Stephenson Caverhill.
Hywel Wyn Edwards.
Frederick Richard Heard.
William Donald Lowe.
Geoffrey Thomas Park.
Vishwa Shanker Shukla.
Eric Lewis Snowdon.
John Thong Ko-Sine.
Colin Osborne Derek Titmus.
Anthony Frederick Twist.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Television Audience in the United Kingdom

By B. P. Emmett

[Read before the Royal Statistical Society, Wednesday, April 18th, 1956.
The President, Professor E. S. Pearson, C.B.E., in the Chair]

Introduction

Most of the data presented in this paper are selected from the considerable quantity collected by the Audience Research Department of the B.B.C. for internal use within the Corporation. The selection includes virtually all the available information concerning changes in the size and nature of the television public of the United Kingdom during the past nine years, but the voluminous material about the viewing habits of the television public (consisting mainly of estimates of the sizes of the audiences of individual television broadcasts) is given only in a summarized form. Information has also been collected in post-war years on a number of other aspects of television, such as the television public's reactions to the programmes they see, the effects of certain series of broadcasts, and so on. This, however, has been omitted entirely, firstly because one section of the paper would be insufficient to do justice to such a wide field, and secondly because at least two important researches are now in progress the results of which should render much of the older material otiose. In no sense therefore is this paper a comprehensive documentation of television and its impact on contemporary society. It is presented in the belief that it contains information not readily obtainable heretofore which may provide the basis for further research into this powerful new medium of communication. A very considerable number of papers and articles concerning television have been published in recent years and the bibliography appended should enable those interested in the field to trace most of them. They include several reports of researches into the wider field of the "effects" of television. Often they are concerned with its effects on behaviour in other respects such as cinema and theatre going, reading, etc.; but even in these limited areas much of the evidence is inconclusive and the effects, if any, in the psychological domain seem only now to be receiving any intensive study. Even though so much research effort has been devoted to the impact of television, there is clearly room for much more, particularly in this country, and it is hoped that those undertaking such research will find basic information about these early years of television's growth useful.

The Size of the Television Public

The television public in the United Kingdom is defined as all persons normally resident in households where there are television receivers. Virtually all the data to be presented refer to those members of the television public 16 years of age and over—the "adult television public". The size of the adult television public may be estimated from the results of the B.B.C.'s daily Survey of Listening and Viewing, for which, each day, some 3,100 interviews are made with persons 16 years of age and over. This number is made up of 400 in each of six B.B.C. Regions and 700 in the seventh (London) Region. The Survey has been running continuously since 1939 but, until 1953, did not cover Northern Ireland. Several alterations in the size of the daily Regional samples have been made during these sixteen years. Until 1952 the number of interviews made each day was 500 in each of the six Regions of Great Britain. Thereafter they stood at 400 in each Region, and it was not until September, 1955, that the sample in the London Region was increased to its present size of 700. Unfortunately, though unavoidably, B.B.C. Regions are quite different from the Registrar General's Standard Regions, being defined by the coverages of the B.B.C. Home Service transmitters. Their boundaries are indicated in Fig. 1. For the purposes of the Survey of Listening and Viewing, each Region is sub-divided into six geographical zones, so defined that the range of variation in the ratio of television sets per head of population in each zone is minimized. This in effect exercises a control over the proportion of each day's informants

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

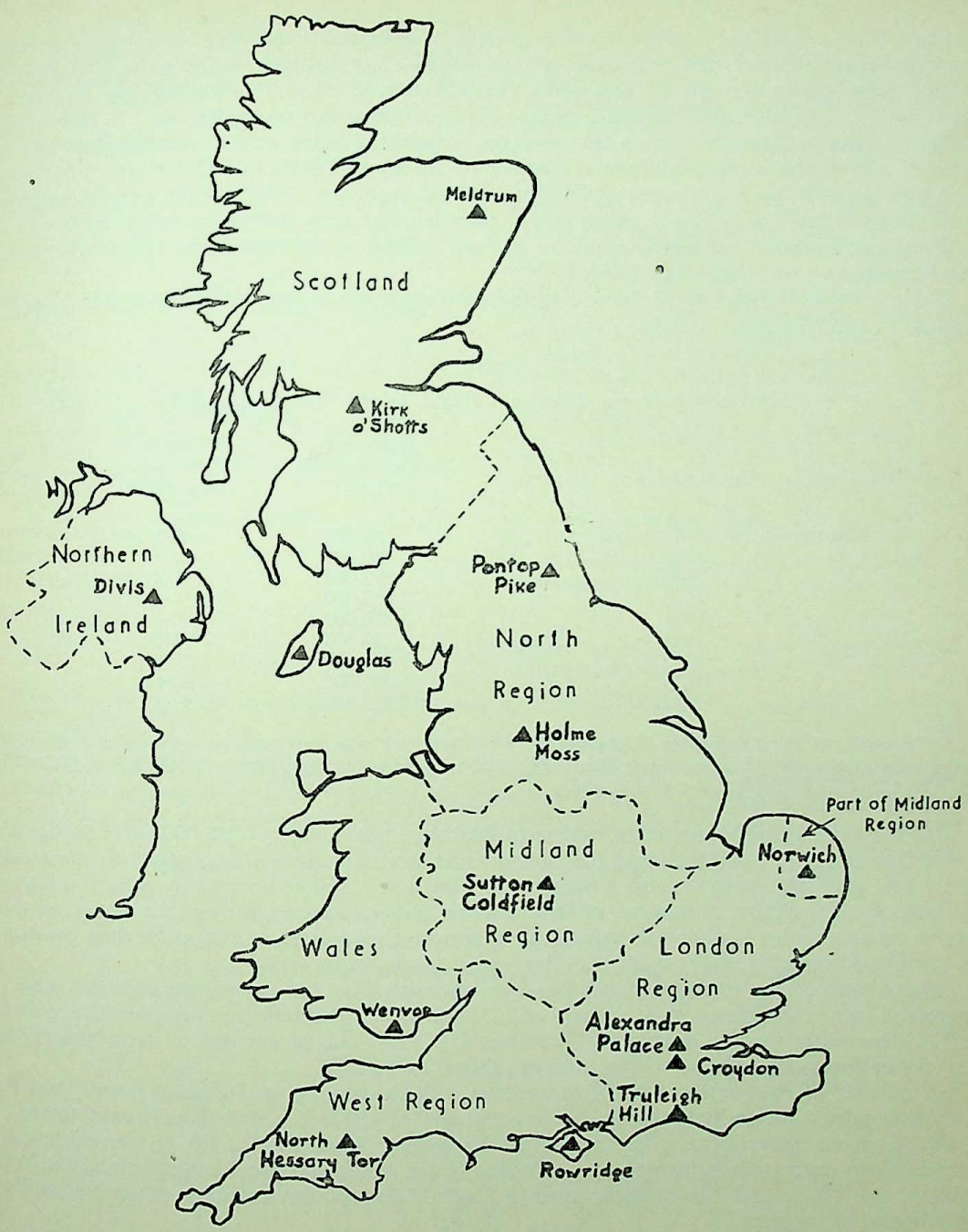


FIG. 1.—Showing boundaries of B.B.C. Regions and locations of the television transmitters operating in 1955.

who are likely to have television. Apart from this, the Survey follows the familiar pattern of quota sampling in this country, with the usual controls of age, sex, social status and employment, though these controls are not, as is frequently the case, inter-related.

The primary object of the interview is to ascertain which programmes, if any, the informant had heard or viewed the day before the interview. In addition, informants are asked whether or not they have television, the precise question being "Have you got television yet?". The

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

presence of the word "yet" at the end of the question may occasion surprise. It is merely a device to overcome any tendency that may exist for informants to feel that they "ought to have television" by giving the opportunity for a "face-saving" answer such as "Well it's on order, but we haven't got it yet". Also surprising, perhaps, is that the question is not more definite. A more exact question might be "Have you got a television set, in working order, at your normal place of residence?". Over the years a substantial number of Survey of Listening and Viewing informants have subsequently been contacted again, for various reasons, and from more precise inquiries made on these later occasions it seems that in very few instances was an incorrect answer given to the simple question actually used in the Survey. There would thus seem to be little reason for changing to a more elaborate form.

Table 1 shows the growth of the television public since the end of 1947 as estimated from the answers to this question.

TABLE 1

The Growth of the Television Public in the United Kingdom

	*Estimated Numbers of Adults with TV**	*TV Public as Proportion of Total Adults* %
At the end of—1947	80,000	0·2
1948	260,000	0·7
1949	670,000	1·8
1950	1,660,000	4·3
1951	3,250,000	8·7
1952	5,250,000	14·0
1953	8,160,000	21·8
1954	11,505,000	30·6
1955	14,925,000	39·8

* To arrive at these estimates of the numbers of viewers it was necessary to estimate also the adult population at the end of each year. This is believed to have risen from about 35,500,000 at the end of 1947 to about 37,600,000 at the end of 1955.

In nine years the television public has increased from a tiny minority of 80,000 (all in the London Region) to nearly 15,000,000 all over the United Kingdom, an increase of almost two-hundred-fold. This rapid growth implies that for most viewers television is still comparatively new. In fact, at the end of 1955, almost one-fourth of the television public had been viewers for not more than twelve months and a further one-half for only two or three years. Little more than one-tenth of them had been viewers for more than five years.

The growth of the television public has been stimulated by the opening during these years of more and more transmitters. At the end of 1955 fourteen B.B.C. television transmitters and one I.T.A. transmitter were in operation. (See Fig. 1.) The dates of opening of these transmitters and the approximate areas served by them are shown on p. 287.

A number of the present B.B.C. transmitters will be working on increased power, and five more low-power transmitters opened, before the end of 1957, and this should make television available in virtually all parts of the United Kingdom. Eight further I.T.A. transmitters are scheduled to open within this period, one covering the Midlands, two Lancashire and the West Riding, one the Scottish Lowlands, another in South Wales and three medium power transmitters to fill some of the gaps.

As each new B.B.C. transmitter opened, the rate of growth of the television public within the area to which it gave good reception* was much the same, the principal exception being Wenvoe. This is shown in Fig. 2, from which it will be seen that the rate of growth of the television public in the *Alexandra Palace* area increased only slowly until the end of 1949 (when it amounted to 7 per cent. of the adult population) and since then has remained roughly constant at about a

* The primary service area of a transmitter is normally taken as being that area enclosed by the 500 micro-volts per metre field strength contour. For the present purpose, the more limited areas within the one milli-volt contours were chosen because not only are they geographically more distinct from one another, but they also generally exclude those areas in which some fading is to be expected.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B.B.C. Transmitters:	*Date of Opening*	*Approximate Coverage Area*
1. Alexandra Palace	2.xi.36	London, M'sex, Herts, Essex, Surrey, Beds, Bucks, Kent (exc. south-east), Sussex (exc. south coast), Berks (east part).
re-opened	7.vi.46	
2. Sutton Coldfield	17.xii.49	Staffs, Warwicks, Northants, Derby, Notts, Salop, Herefs, Worcs, Oxon (north), Gloucs (north).
3. Holme Moss	12.x.51	Lancs, Yorks, Cheshire, Lincs, Cumbs (south), Flint, Denbigh, Caernarvon (east).
4. Kirk o'Shotts	14.iii.52	Ayr, Lanark, Peebles, E. Lothian, Midlothian, W. Lothian, Renfrew, Dumbarton, Stirling, Clackmannan, Fife, Perth (south).
5. Wenvoe	15.viii.52	Glams, Carmarthen (south-east), Brecknock (south), Monmouth, Gloucs (south), Wilts, Somerset, Dorset, Devon (north).
6. Pontop Pike	1.v.53	Durham, most of Northumberland.
7. Glencairn	1.v.53	
(now superseded by Divis)	21.vii.55)	Antrim, Down.
8. Truleigh Hill	9.v.53	Brighton and immediate environs.
9. Douglas	20.xii.53	Isle of Man.
10. Rowridge	12.xi.54	Hants, Isle of Wight, Sussex (west).
11. Redmoss	14.xii.54	
(now superseded by Meldrum)	12.x.55)	Aberdeen, Kincardine, Banff.
12. North Hessary Tor	17.xii.54	Devon, Cornwall.
13. Norwich	1.ii.55	Norwich and immediate environs.
14. Les Platons	3.x.55	The Channel Islands.
I.T.A. Transmitter:		
Croydon	22.ix.55	London, M'sex, Surrey, parts of Beds, Bucks, Herts, Essex, Kent, Berks.

These transmitters make available a television service to 94 per cent. of the total population of the United Kingdom.

further 7 per cent. per year, reaching 48 per cent. by the end of 1955. The *Sutton Coldfield* transmitter opened at the end of 1949. The rate of growth in the area it serves has been slightly more rapid than in the Alexandra Palace area, particularly in 1950 and 1951, so that by the end of 1954

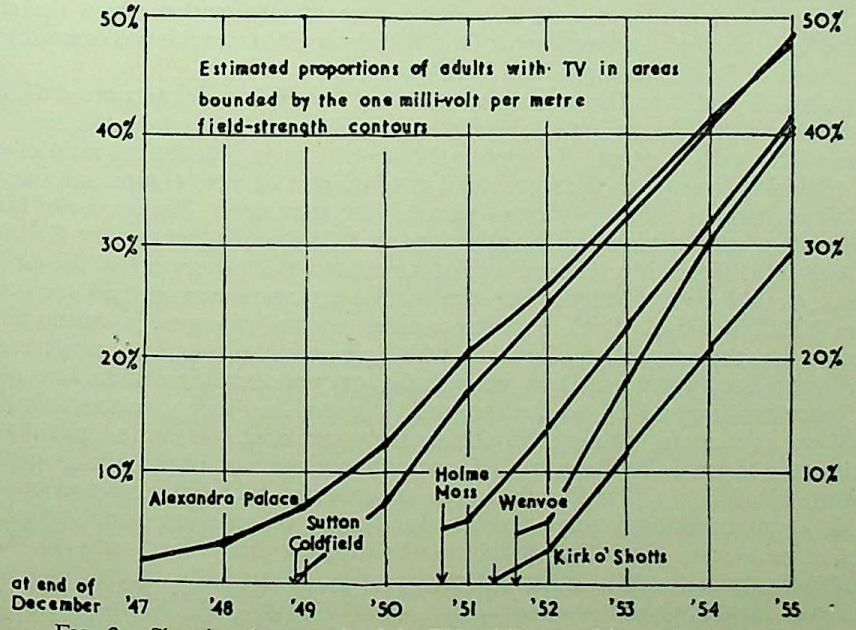


FIG. 2.—Showing the growth of the TV public served by each of the five principal transmitters.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the two areas had roughly equal proportions of adults with television. By the end of 1955 the Sutton Coldfield proportion even exceeded that in the Alexandra Palace area. When the *Holme Moss* transmitter opened at the end of 1951, the area it served already contained some viewers (7 per cent. of its adult population) receiving Sutton Coldfield, but from 1952 onwards the rate of growth has been very similar to that of the Sutton Coldfield area. The next transmitter to open was *Kirk O'Shotts*, in mid-1952, and there also the expansion of the television public has been fairly steady at about 8 per cent. of the adult population each year. The last of the present high-power transmitters to open was *Wenvoe*, in the autumn of 1952. As with the Holme Moss area, there were already some viewers (5 per cent.) in the area now served by Wenvoe before the transmitter opened. In its first two years of operation, 1953 and 1954, the television public there grew more rapidly than it did elsewhere, the rate being some 11 per cent. per year. During 1955 the rate of growth has been, if anything, even greater, so that by the end of the year its public amounted to 41 per cent., nearly as much as in the Alexandra Palace and Sutton Coldfield areas. There is no obvious reason why the Wenvoe area should prove an exception and it remains to be seen whether the rapid growth there will be maintained. In the areas served by the four low-power transmitters opened during 1953, the rates of growth of the television public have all been "normal", i.e. between 6 per cent. and 8 per cent. of the adult population per year.

The "density of TV"—a conveniently brief term for "the numbers of adults with television as a percentage of all adults"—varies considerably even within these areas of good reception. The Survey of Listening and Viewing cannot give adequate estimates of the densities in respect of geographical units smaller than the zones into which B.B.C. Regions are divided, but there is another source which can throw light on the geographical variations. The General Post Office issues monthly statements of the numbers of sound and television licences current in the United Kingdom, derived by the Accountant General's Department from licence counterfoils. Every third month a more extensive statement is issued showing the numbers of current licences in each Head Post Office District of the United Kingdom, these figures being derived from returns made by Head Postmasters. Licences may be taken out at nearly all Post Offices and the counterfoils are subsequently sent to the Head Post Offices of the districts in which the licencees live. There is inevitably a discrepancy between the total of the licences shown on the Head Postmasters' returns and the total given by the accounting Department—the latter being presumably the more exact—since the process of transmitting counterfoils from one post office to another, and the filing of them on receipt, must make an exact count on any one day virtually impossible. However, the Head Postmasters' returns remain the only source of information about the geographical distribution of licences, and the magnitude of the discrepancy between the two estimates is generally not large enough to be of much practical importance.

The populations and even the boundaries of Head Post Office districts are difficult to determine, and without such information it is clearly impossible to make precise estimates of the density of television within each of them. It appears, however, that in practice the ratio of the number of current television licences to the total number of current licences (sound *and* television) may provide a fair estimate of the proportion of persons with television. The Survey of Listening and Viewing estimate of the television public at the end of 1954 was 30·6 per cent. of the adult population. At that time the television licences current amounted to 30·2 per cent. of the total. Almost equally exact was the correspondence between the Survey estimates and the "licence ratios" for each of the B.B.C. Regions. Such a correspondence implies, of course, a relationship between several factors, e.g. the proportion of persons who have neither a sound nor a television receiver, the proportions of sound and television receivers that are unlicensed, the mean sizes of the households of licence evaders and others, and so on. The fact that such a relationship appeared to hold in 1954 was presumably nothing more than a happy accident. It even seemed that the simple ratio of television to total licences provided a satisfactory estimate of the density of television in any individual Head Post Office District. The adequacy of such estimates was investigated by analysing, in terms of administrative districts, the Survey of Listening and Viewing results for four weeks in mid-1954. Each Survey interviewer works within a given administrative district and most of them make quotas of fifteen interviews a day. Thus in the course of a week a sample of 105 adults is interviewed by each of them. The four weeks studied yielded 683 samples of 105 interviews each, made in all kinds of administrative districts from large cities to rural districts. The proportions of informants with television derived from these small samples were compared

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

with the ratio estimates for the most closely corresponding Head Post Office Districts. Table 2 gives the bivariate frequency distribution of these 683 sets of results.

TABLE 2

Percentage of Survey of Listening and Viewing Informants with TV

Ratio of Television to Total Licences per cent.	0–3	4–7	8–11	12–15	16–19	20–23	24–27	28–31	32–35	36–39	40–43	44–47	48–51	52 and over	Total
0–3	34	12	3	2	1	—	—	—	—	—	—	—	—	—	52
4–7	6	13	7	2	1	3	—	—	—	—	—	—	—	—	32
8–11	—	9	2	6	5	6	3	2	1	—	—	—	—	—	34
12–15	—	8	8	12	8	3	2	—	—	—	—	—	—	—	41
16–19	2	5	13	26	25	18	8	4	1	—	—	—	—	—	102
20–23	—	—	2	12	18	18	10	6	3	3	—	—	—	—	72
24–27	—	—	2	6	7	14	13	10	8	3	3	—	—	—	66
28–31	—	—	—	2	4	12	23	20	16	7	6	2	—	—	92
32–35	—	—	—	—	2	3	10	18	9	10	5	1	1	—	59
36–39	—	—	—	—	—	1	3	17	23	20	12	11	7	—	94
40–43	—	—	—	—	—	—	—	—	3	4	3	4	0	1	15
44–47	—	—	—	—	—	—	—	1	4	3	3	3	2	4	20
48–51	—	—	—	—	—	—	—	—	1	—	—	1	1	1	4
52 and over	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
Total	42	47	37	68	71	78	72	78	69	50	32	22	11	6	683

An exact correspondence between these two sets of figures could not be expected since not only are the Survey of Listening and Viewing samples in the administrative districts small but these districts may well be quite different in extent from the Head Post Office districts. A number of the biggest discrepancies shown above can almost certainly be ascribed to differences in the areas to which the two estimates relate. Nevertheless, there is sufficient indication that, not only for the United Kingdom and each B.B.C. Region, but also for individual Head Post Office Districts, the simple licence ratio estimates are likely to serve at least as rough and ready guides to the true densities of television.

It is not at all certain that this simple and convenient measure will continue to provide adequate estimates of the geographical densities. In fact, throughout 1955 a tendency became apparent for the licence ratios to fall consistently short of the Survey of Listening and Viewing estimates. Fortunately, however, the disparities so far are comparatively slight.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

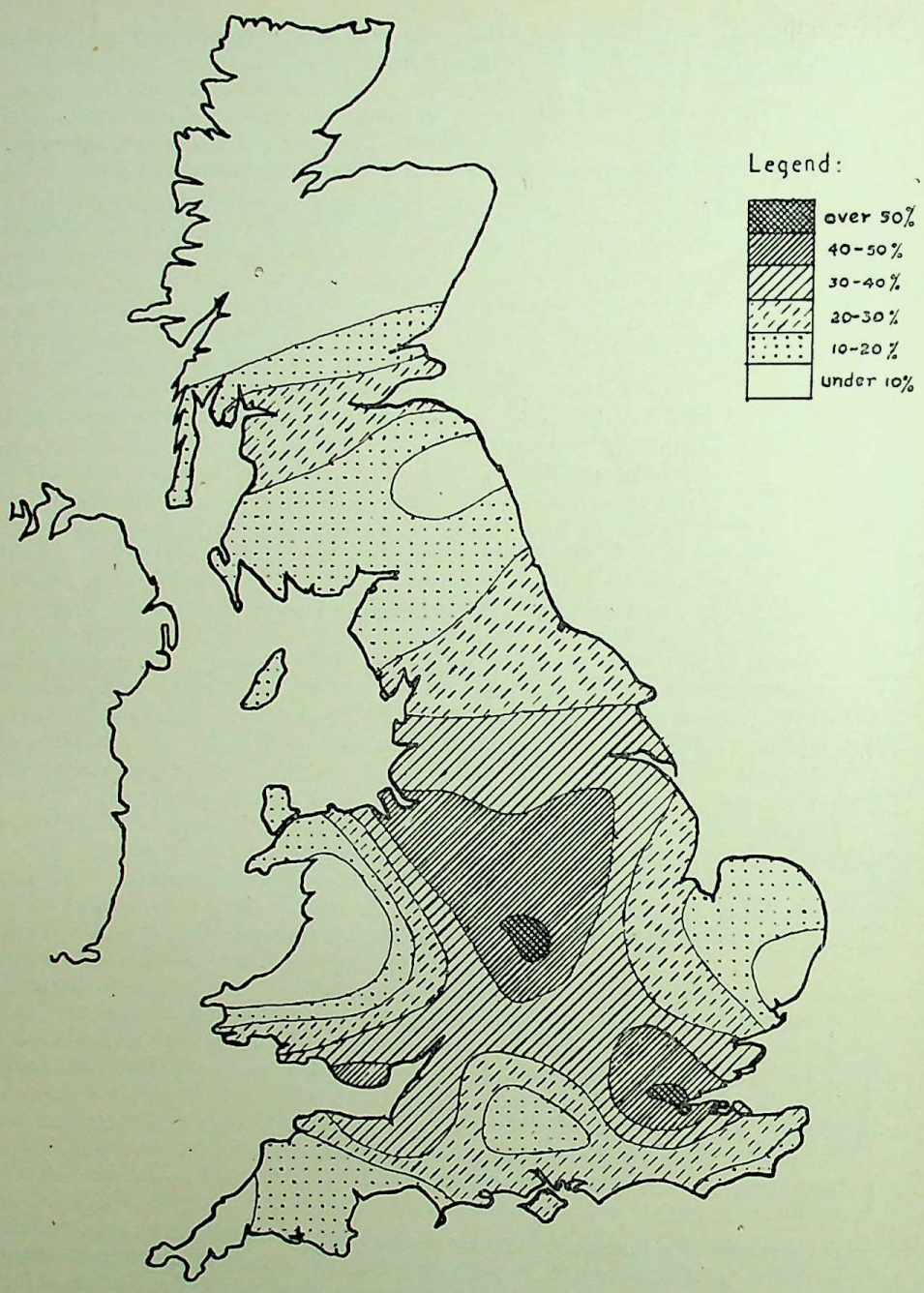


FIG. 3.—Showing the density of television in the United Kingdom at the end of September, 1955.

Fig. 3, showing the density of television in the United Kingdom, is based on the licence ratios for the end of September, 1955. It will be seen that the highest density of television was to be found in the northern, eastern and south-eastern suburbs of London, from Edgware to Dartford, where well over 50 per cent. of the population then had television. The density was almost equally high (just over 50 per cent.) in the Tamworth, Coventry, Nuneaton, area of the Midlands. Between 40 per cent. and 50 per cent. of the population had TV in that part of the London Region stretching

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

from Hitchin in the north to Croydon in the south and from Slough in the west almost to Southend in the east. A similar density was to be found in a much larger area of the Midlands with boundaries running on the east from Rugby to Scunthorpe, on the north from Scunthorpe to Liverpool, and on the west from the estuary of the Dee to Redditch. A third area with a density of over 40 per cent. was in Glamorganshire, from Bridgend to Cardiff. Densities of between 20 per cent. and 40 per cent. were found in many parts of the country, notably in the rest of the Home Counties, in the West Riding and Tyneside conurbations, in Lancashire north of Manchester and in the Lowlands of Scotland. In only six areas, all with comparatively small populations, did the density fall below 10 per cent., viz. the Highlands of Scotland, the Border country around Kelso and Hawick, in most of Northern Ireland, in central and west Wales, in Cornwall, and in parts of East Anglia.

In general the variations in density tend to parallel field strength contours, suggesting that, as might be expected, the quality of reception is a potent factor in causing them. But the parallelism is not so close as to indicate that the reception quality is the *only* factor at work. For instance, it is in Woodford and Romford that the highest densities have been found, but the reception there should be no better than in most north London suburbs. Indeed, though field strengths decrease with distance from the transmitters, reception should nevertheless be perfectly satisfactory over a considerable area. Several hypotheses can be formulated to explain some, if not all, of the variations. Two possible factors are (*a*) the general economic levels of the districts and (*b*) the extent to which their inhabitants form closely-knit societies. It will be shown later that possession of television is associated with income level, and in a closely-knit society that powerful motive of "keeping up with the Joneses" might perhaps contribute towards a relatively more rapid spread of television. Intensive sales campaigns by retailers in some localities could also have made an important contribution to the variations in density, and so could the proportions of households with children. (The association between the presence of children in the home and the possession of television is an interesting topic that will be dealt with later.) No doubt other contributory causes could be suggested. It is important to note, however, that all these factors are related to the rate of acquisition of television which, except where reception is impossible, is never very low. Even where reception is poor, the television public is expanding steadily at a rate of some 3 per cent. per year, i.e. about half as fast as the "normal" 6 per cent. in the good reception areas.

The expansion occurring during the past few years leaves little doubt as to the progress in the next few years. In some parts of the United Kingdom persons without television are already a minority and there seems every reason to believe that this will be true of most parts of the country in the comparatively near future since there is as yet no sign that the rate of increase is decreasing even in those areas where the television public is now in the majority. In fact, from the present state of affairs it appears certain that, even as soon as 1960, more than three-quarters of the population will have television.

A new element was introduced into British television with the opening, in September, 1955, of the London I.T.A. transmitter. Before they could receive the new service, most viewers needed either to buy a new set or to have their present sets adapted, and nearly all of them needed a new aerial. Since the service started, the growth of the public for Independent television has been measured by the Survey of Listening and Viewing. The proportions of informants in the London Region claiming to be able to receive I.T.A. transmission are given in Table 3.

During the first four weeks of independent television, about one-tenth of the adult population of the London Region could receive the new programme on their sets, the proportion increasing

TABLE 3

The I.T.A. Public in the London Region

		As Proportion of All Adults %	*As Proportion of Adults with TV* %
1955	Weeks 39–42 . . .	9·5	22·8
	43–46 . . .	12·7	28·9
	47–50 . . .	14·5	32·6
	51– 2 . . .	16·2	36·6

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

steadily as more sets were converted, or new sets bought, to about one-sixth by the end of the year, i.e. about a million and a half adults.

The geographical variations in density of the I.T.A. public cannot be estimated from the licence data by the ratio method used for the television public as a whole. It will be recalled, however, that each B.B.C. Region is divided into six geographical zones so defined that the variations in density of television within each is a minimum. It is on the basis of the licence ratios that these zones are defined.

At the end of November, 1955, the densities of the I.T.A. public in the six London Regional zones were found to be as follows:

TABLE 4

The I.T.A. Public at the End of November, 1955

	As Proportion of All Adults %	*As Proportion of TV Public* %
In—Zone 1	17·4	31·7
Zone 2	19·6	42·2
Zone 3	15·1	29·4
Zone 4	6·8	18·8
Zone 5	3·7	14·0
Zone 6	—	—
All Zones	14·4	31·7

The coverage area of the I.T.A. Croydon transmitter consists of Zones 1 and 2 (together comprising Greater London), most of Zone 3 (a semi-circular band north of London) and part of Zone 4 [another semi-circular band south of London, but reaching as far as the south coast]. It seems certain, therefore, that by the end of November, 1955, of the 8½ million adults in the coverage area about 17 per cent. (1,400,000) were able to receive the I.T.A. programmes.

Before turning to other aspects of television in the United Kingdom than the size of its public, it may be of interest to review the position in other countries. A U.N.E.S.C.O. report in 1953 gave, *inter alia*, such details as were available of all television transmitters in the world known to be operating then, and of the numbers of television receivers in use in each country. A supplement was issued in 1955 bringing this information up to date. Table 5 has been derived from these two reports. Besides showing for each country the numbers of transmitters and receivers in 1955, it also gives (as did the 1953 U.N.E.S.C.O. report) rough estimates of the numbers of inhabitants per television receiver.

Many of these data are somewhat speculative, particularly the numbers of receivers shown. The "numbers of inhabitants per receiver", even if the figures were reliable, would not be good indices of the penetration of television amongst the populations of these countries but, in the absence of any information about family size or the numbers of persons normally having access to each receiver, nothing more satisfactory can be devised. At least it seems safe to say that, whatever the mean family sizes, the countries with the highest densities of television at the end of 1954 must have been, in order, the United States, the United Kingdom, Canada, Puerto Rico and Cuba; and that in no other country did the density of television approach the level even of the last of these.

It is important to realize that the situation is changing very rapidly indeed, and that the picture as revealed by Table 5 cannot therefore be taken to represent the present position. Both the numbers of transmitters and of receivers are increasing—very rapidly in some countries and very slowly in others. For instance the first television transmission in Canada did not take place until September, 1952, but the growth there has been extremely rapid. In fact, recent information from the Canadian Broadcasting Company shows that by the end of 1955 the density of television there, amounting to some 46 per cent. of the population, considerably exceeded that in the United Kingdom. In France, on the other hand, transmissions were given as early as 1935, but, even since the war, the number of viewers has remained comparatively small. To quote the U.N.E.S.C.O report: "The significant aspect of the present (1953) phase of (television's) spread is that it is beginning to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 5

The Number of Television Transmitters and Receivers in Various Countries

	Number of TV Transmitters Operating during 1955	*Number of TV Receivers at the End of* 1954	*Persons per TV Receiver*
Alaska	4	12,000*	
Argentina	1	30,000	850
Australia	Experimental		
Austria	,,		
Belgium	4	15–40,000	
Brazil	5	120,000†	380
Bulgaria	1		
Canada	29	1,100,000	15
Columbia	1‡		
Cuba	9	135,000	50
Czechoslovakia	1 or 2		
Denmark	1	2,390	1,750
Dominican Republic	1	5,000	450
France	10	209,000	210
Germany, East	4	4,000	
,, West	23	160,000	300
Hawaii	3	62,000	80
Hungary	Experimental		
Italy	9	82,000	480
Japan	5	33,797	3,500
Luxembourg	1		
Mexico	14	100,000	300
Monaco	1		
Morocco	2	2,500	
Netherlands	1	15,000	700
Norway	Experimental		
Philippines	1	1,513	
Poland	Experimental		
Puerto Rico	2	65,000	35
Saar	2		
Spain	Experimental		
Sweden	,,		
Switzerland	5 (mainly experimental)	4,930	1,000
U.S.S.R.	10§	700,000	270
U.K.	14	4,600,000	11
U.S.	413	35,000,000	4·5
Venezuela	3	30,000	170
Yugoslavia	Experimental		

* Fairbanks City only. Numbers elsewhere not known.
† Autumn, 1954.
‡ Government transmitter for mass education.
§ Including amateur stations.

establish itself in industrially much less advanced countries, and thereby is bringing about a fundamental change in the world communication picture"—whether for good or ill who would be rash enough to say?

THE NATURE OF THE TELEVISION PUBLIC

The nature, as well as the size and geographical distribution, of the television public has been steadily changing since the war. The Survey of Listening and Viewing can give little information about the nature of the television public since no "personal details" questions are asked and hence only the classifications made for the quota controls are available for analysis. This somewhat meagre supply can, however, be amplified from the results of three special television inquiries, the first made in the spring of 1948, the second in the autumn of 1950 and the third in the spring of 1954. The 1948 inquiry studied a random sample of television-owning families in Greater

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

London and a "control group" of next-door-neighbours *without* television. The 1950 survey did likewise, but also included informants from the Sutton Coldfield coverage area.

The sample for the 1954 inquiry was rather more unorthodox, being a sub-sample of informants in an earlier inquiry. In October, 1953, B.B.C. Interviewers had made brief interviews at some 9,000 households in the United Kingdom. The primary sampling units were 99 administrative districts, not chosen at random but because reliable interviewers were available in them. The districts were, however, stratified geographically by B.B.C. Regions and, within Regions, by size of population. The second stage consisted of a random selection of three polling districts in each administrative district from comprehensive lists stratified by the "juror count" method and with probability proportional to the number of electors. Finally, clusters of 30 addresses in each polling district were chosen, taking every fifth address from a randomly selected starting point. The normal Survey of Listening and Viewing question "Have you got television yet?" was asked and an enumeration made of all members of the household.

For the 1954 television inquiry, four months later, 83 of the 99 interviewers who had taken part in this survey were available. They called back on all households in their administrative districts which, in October, 1953, had been found to have television, and at one in four of all households without television but with sound listening facilities. About 20 per cent. of the households had been found to have television so, by taking one in four of the remainder, the television and sound samples became roughly equal in numbers. At each household they interviewed one adult chosen at random from amongst all the adults in the household, and, to convert to a sample of individuals, the punched cards were reproduced according to the number of adults in the household. Furthermore, special provision was made for those households in which television had been acquired since October, 1953, the appropriate cards being reproduced four times and added to the television pack. About 7,100 successful interviews had been made in the 83 administrative districts, including 1,437 at households with television and 5,260 with sound radio but not television (of which one-fourth, or 1,315, were called on again). 1,322 successful interviews were made in television households and 1,171 in non-television households, 92 per cent. and 89 per cent., respectively, of the totals. Despite the unorthodox nature of the method, the samples achieved exhibited no apparent bias when compared in a number of respects with either official data or the results of other surveys.

Social status. In all three special television inquiries, questions were asked about the income of the head of the household. The results were broadly consistent, i.e. the income distribution of the comparatively small number of viewers who, in 1954, stated they had been viewers since 1950 agreed closely with the income distribution of the 1950 sample; and the income distribution of the even smaller number of 1954 informants who had been viewers since 1948 agreed with that of the 1948 sample. Checks on the intervening years were made by comparing the 1954 results with those obtained in the annual Hulton Readership Surveys which, since 1949, have given estimates of the income distribution of the television public. Once again the agreement was sufficiently close—despite differences in the income groupings used—to suggest that the 1954 inquiry results are fairly accurate estimates.

TABLE 6

Changes in the Income Distribution of the Television Public

	Class I (Top 12%)		*Class II (Next 20%)*		*Class III (Bottom 68%)*		*All Adults*	
	(,000)	%	(,000)	%	(,000)	%	(,000)	
At end of 1947 . . .	42,	48	20,	27	20,	25	80,	100
New in 1948 . . .	65,	36	65,	36	50,	28	180,	100
,, 1949 . . .	120,	29	135,	33	155,	38	410,	100
,, 1950 . . .	230,	23	315,	32	445,	45	990,	100
,, 1951 . . .	335,	21	545,	33	710,	46	1,590,	100
,, 1952 . . .	380,	19	645,	32	995,	49	2,020,	100
,, 1953 . . .	360,	12	625,	21	1,925,	66	2,910,	100
,, 1954 . . .	285,	9	505,	16	2,535,	75	3,325,	100
At end of 1954 . . .	1,815,	16	2,855,	25	6,835,	59	11,505,	100

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In Table 6 the population is divided into three strata, thus: Class I, the better off, 12 per cent. (i.e. those families in which the head's income, in 1954 terms, would be over £15 a week), Class II, the next 20 per cent. (£10 to £15 a week) and Class III, the bottom 68 per cent. (less than £10 a week). It shows the approximate numbers of adults in each class who had television sets at the end of 1947, the numbers who acquired them year by year, and finally the numbers as they stood at the end of 1954. This latter estimate [for the end of 1954] is an extrapolation from the 1954 inquiry results, guided by the 1955 Hulton Readership Survey.

From this Table it will be seen that at the end of 1947 the television public was very far from being a "cross-section" of the population; nearly half (48 per cent.) came from Class I, which constitutes only 12 per cent. of the population, whereas only a quarter came from Class III, which constitutes 68 per cent. In fact, at that time, nine out of every 1,000 Class I persons had television compared with less than one out of every 1,000 Class III persons.

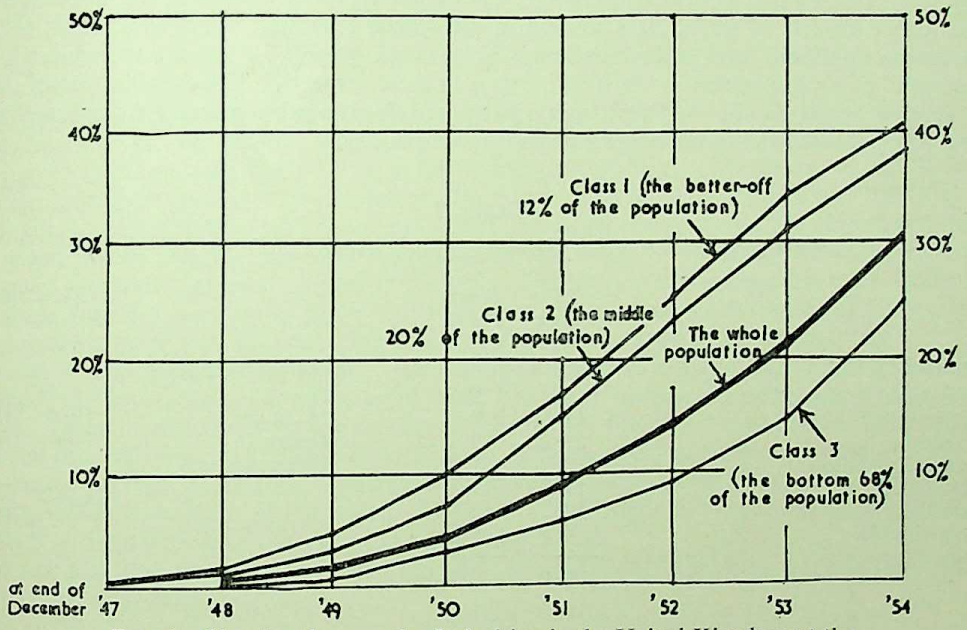


FIG. 4.—Showing the growth of television in the United Kingdom at the end of September, 1955.

In the following years more and more people of each class became viewers, but each annual intake was less "top-heavy" than the last. In 1952, for example, Class I's contribution was down to 19 per cent. and Class III's up to nearly half. This was far less top-heavy than it had been five years earlier, but it still represented a recruitment of 85 out of every 1,000 Class I people as against only 39 out of every 1,000 of the Class III population. By the end of 1953 there were signs that the top-heavy tendency would disappear altogether. This was confirmed during 1954, when it seems that Classes I and II were contributing *less*, and Class III *more*, than their proportionate shares to the new entrants. The substantive position at the end of 1954 can be seen in the bottom line of the Table. By then the television public, which numbered over eleven and a half millions as compared with only eighty thousand eight years earlier, was coming much closer to being a cross-section of the population, Class I contributing 16 per cent., Class II 25 per cent. and Class III 59 per cent.

Fig. 4 shows the rate at which each Class has been acquiring television. The rate of expansion in Classes I and II was more or less constant during 1951, 1952 and 1953 at about 8 per cent. per year, but fell slightly during 1954. The rate of expansion in Class III was still increasing at the end of 1954, the increase during that year amounting to 10 per cent. of the adults in that Class.

No information is available which would enable this analysis by income to be carried further forward in time, but an analysis of material from the Survey of Listening and Viewing can bring

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the story more up to date. Social class is one of the quota controls in this Survey, interviewers making 5 per cent. of their interviews with persons in the Upper Middle Class, 20 per cent. with persons in the Lower Middle Class and 75 per cent. with Working Class people. A study of this material tells very much the same story as that set out in Table 6—that from being predominantly Upper and Lower Middle Class the television public is every year coming nearer to being a cross-section of the population. According to these results, the position at the end of 1955 was that the television public included 40 per cent. of the entire adult population, among which were 48 per cent. of the Upper Middle Class, 46·5 per cent. of the Lower Middle Class and nearly 38 per cent. of the Working Class. Moreover, it was clear that recruitment from the two Middle Classes was proceeding more slowly in 1955 than in previous years. Some 8 per cent. of them became viewers in 1955, as against nearly 9 per cent. in each of 1953 and 1954. In the Working Class the rate of recruitment in 1955 was about the same as in the Middle Classes, but whereas the Middle Class rate tended to decrease slightly, that of the Working Class was still increasing.

Two other aspects of social class studied in the special television inquiries of 1948, 1950 and 1954 were occupational level and educational background. Table 7 contrasts the adult television and "sound" publics, as given in the 1954 inquiry, in these respects. The classification of occupational level is based on the skill and training required for the job; married women and retired persons being classified in terms of their previous occupations.

TABLE 7

Comparison of the Adult Television and Sound Public in March, 1954

	TV Public %	*Sound Public* %
Occupational level:		
Professional, technical and executive	8·4	5·5
Highly skilled	11·5	8·6
Skilled	29·2	23·1
Moderately skilled	19·7	21·2
Semi-skilled	17·0	19·8
Unskilled	10·2	16·0
Unclassifiable	4·0	5·8
	100·0	100·0

	TV Public %	*Sound Public* %
School leaving age:		
Over 18	5·4	4·4
16–18 yrs.	19·4	17·6
Under 16	74·4	77·7
Unknown	0·8	0·3
	100·0	100·0

Both these comparisons show the television public, at the beginning of 1954, to have been slightly "superior" to the sound public; its mean occupational level was higher and rather more of the television than of the sound public had had a full secondary or higher education. But, as the previous paragraphs have suggested, the economic difference between viewers and non-viewers is diminishing and almost certainly the same is true of occupational and educational levels. The time is coming, if indeed it has not already come, when there will be little to distinguish the television public from a cross-section of the population in these respects.

At this point it is pertinent to recall a finding of the 1948 television inquiry (which was confirmed by that of 1950). While it was true that the larger the bread-winner's income the greater was the likelihood that a family would have a television set, given middle class families of roughly equal economic status but of unequal educational level, those of the lower educational level would be likely to be the ones which bought television sets first. The 1954 inquiry showed this still to be so. This will be seen from Table 8, in which the educational structure of the television public in each of four economic strata is contrasted with that of the sound public in these strata.

It will be seen that the educational level of viewers in the highest economic stratum (head of household's income over £25 per week) is distinctly lower than that of listeners in the same stratum —25 per cent. of these viewers, as compared with 40 per cent. of these listeners, are of the top educational level. The educational level of viewers in the second economic stratum (£15–£25 per week) is also below that of listeners in the same stratum. The same is true of the third economic stratum (£10–£15 per week), but in the fourth, the educational levels of viewers and listeners appear to be much the same.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 8

Income Class and Educational Background of the Television and Sound Publics in Spring 1954

	Income of Head of Household							
	Over £25 per week		*£15–£25 per week*		*£10–£15 per week*		*Under £10 per week*	
	TV %	*Sd* %	*TV* %	*Sd* %	*TV* %	*Sd* %	*TV* %	*Sd* %
Full-time education finished at:								
Over 18 years	25	40	10	16	5	7	3	2
16–18 years	59	44	38	50	22	28	10	11
Under 16 years	13	16	51	32	73	64	87	87
Unknown	3	—	1	2	—	1	—	—
	100	100	100	100	100	100	100	100

It is important that this phenomenon should be seen in perspective. In the first place it applies only to the upper income groups, which accounted for less than half the television public at the beginning of 1954, and are likely to be a diminishing proportion. Moreover, it would not be true to say that, even among them, the lower their educational level the more likely they are to be viewers, for educational level and economic status are closely related and the latter seems to be the far more potent factor in bringing television into the home. What can be said, however, is that among the middle classes the higher the level of education the greater the sales resistance—though this does not necessarily mean that television appeals less to the better educated than to others; it could mean that to them other forms of expenditure appear more important.

Household size and age composition. The average number of persons per television household was found, at the beginning of 1954, to be 3·42 (the 1950 television inquiry showed it then to be 3·43). These figures are above the average for the rest of the population (3·17). The principal reason for this difference is that among television families, the "one-person household" is very rare—about 2 per cent. of the total—whereas such households constitute 10 per cent. to 12 per cent. of the rest of the households of the United Kingdom. This is hardly surprising when it is recalled that many of them consist of old-age pensioners living alone.

The age distribution of the television public is also different in certain respects from that of the rest of the population.

TABLE 9

	Mean Number of Persons per Household	
	With TV	*With Sound Radio but Not TV*
Age:		
0– 4 years	0·20	0·22
5–15 ,,	0·70	0·50
16–25 ,,	0·33	0·33
26–45 ,,	1·18	0·88
46–65 ,,	0·78	0·91
over 65 ,,	0·23	0·33
	3·42	3·17

There has been little change in the age composition of the television public as it has grown over the years. The 1950 inquiry and an analysis of the yearly influxes of new entrants since then as given by the 1954 inquiry showed the only significant change to be a diminishing proportion of 0–4-year-olds, and a corresponding increase of 5–15-year-olds, consequent upon the decline in the birth rate from its exceptionally high level just after the war.

The differences shown in Table 9 in the composition of television and sound households have also remained constant over the years, the 1954 inquiry results being very similar to those yielded

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

by the earlier inquiries. That there were a third more 26–45-year-olds in television than in sound families suggests that the presence in a family of more than one working adult increases the likelihood of its acquiring a television set. The comparative scarcity of the elderly amongst television families arises, no doubt, from the same cause as does the scarcity of one-person households, viz. that few old-age pensioners living alone are likely to have television. That television households contained nearly half as many again 5–15-year-olds as did sound households suggests that the presence of these children makes the acquisition of a television set more likely. It was as part of a special inquiry amongst 5–20-year-olds that the 9,000 interviews were made that formed the basis of the 1954 television inquiry sample. It was then found that amongst households that contained no minors (5–20-year-olds), 19 per cent. had a television set, whereas among the families that included one or more minors, 28 per cent. had television. (Television sets were found in 28 per cent. of one-minor households, in 33 per cent. of two-minor households, in 24 per cent. of three-minor households and in 20 per cent. of households with four or more minors. It seems, therefore, that a more precise conclusion might be that the presence of *up to three* 5–15-year-olds in a family makes the acquisition of a television set more likely).

Viewing facilities. The 1948, 1950 and 1954 television inquiries all gathered information about the screen sizes in use at the time. Table 10 sets out the results.

TABLE 10

Size of Screen of Television Set

	Spring 1954 %	*Autumn* 1950 %	*Spring* 1948 %
15 in. × 12 in. or larger . . .	22	—	—
12 ,, × 10 ,, ...	53	18	11
10 ,, × 8 ,, ...	17	34	29
8 ,, × 6 ,, ...	7	38	51
Other sizes	—	9	7
Unknown	1	1	2
	100	100	100

There is reason to believe that people are not always truthful when they answer questions about screen size—there is a tendency to exaggerate, presumably for reasons of prestige. In the 1954 figures, for example, it is probable that the proportion of large size screens is inflated and the proportion with 8 in. × 6 in. screens too small. Nevertheless there seems no doubt that there is a trend towards larger sizes; whatever the exaggeration it seems certain that the 8 in. × 6 in. from being easily the most common size in 1948 was relatively rare by 1954, its place being taken by the 12 in. × 10 in. By the end of 1955 the trend towards larger screens had no doubt progressed even further. All three inquiries revealed that virtually everyone with a television set also has facilities for listening to sound broadcasting. In the spring of 1954, 3 per cent. of viewers received sound broadcasts by radio relay, 96 per cent. had an ordinary sound receiver, and only 1 per cent. had no listening facilities.

THE VIEWING OF THE TELEVISION PUBLIC

It will be recalled that, in the daily Survey of Listening and Viewing informants are asked which programmes they saw or heard on the day preceding the interview. The interviewer assists recall by helping the informant to reconstruct his activities and by reminding him which programmes were on the air. The items heard or viewed are recorded on "log sheets" which contain comprehensive lists of all programmes broadcast. After the sound radio items have been "logged" informants are asked whether or not they have television and the TV items are logged. The interview ends with a general question about satisfaction with current programmes, both sound and television, and a request for the informant's name, address and occupation.

These interviews yield estimates of the size of audience of each item. The material is far too voluminous to give in detail, but Table 11 presents it in summarized form, showing averages of the audiences for items broadcast in the daytime (i.e. before 6 p.m.) and in the evening (6 p.m. to 11 p.m.).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 11

Average Audiences for Television amongst the Adult Population

	Average Daytime TV Audience %	*Average Evening TV Audience* %
Year ending:		
December 31st, 1952	1·6	5·8
March 31st, 1953	1·7	6·5
June 30th, ,,	1.8	7·2
September 30th, ,,	2·0	7·9
December 31st, ,,	2·1	8·7
March 31st, 1954	2·2	9·3
June 30th, ,,	2·3	10·2
September 30th, ,,	2·4	10·6
December 31st, ,,	2·8	11·4
March 31st, 1955	2·8	12·2
June 30th, ,,	2·8	12·9
September 30th, ,,	2·8	13·1
December 31st, ,,	3·0	13·7

These figures show that the average audience for evening television broadcasts rose from just over two million in 1951 to over five million three years later. Since this increase closely paralleled the rise in the number of television sets in use, it may be deduced that the extent to which people with television receivers used them must have remained pretty constant. Evidence on this is available for the past three years and confirms this deduction. It is set out in Table 12 which shows the averages of the audiences for daytime and evening items *amongst the adult television public* (figures are not available for 1952 or earlier).

TABLE 12

Average Audiences amongst the Adult Television Public

	Average Daytime TV Audience %	*Average Evening TV Audience* %
Year ending:		
December 31st, 1953	9·9	39·0
March 31st, 1954	9·8	39·9
June 30th, ,,	9·9	40·6
September 30th, ,,	9·7	40·6
December 31st, ,,	9·6	40·6
March 31st, 1955	9·7	40·6
June 30th, ,,	9·0	39·6
September 30th, ,,	8·4	37·9
December 31st, ,,	8·1	36·6

The constancy of the viewing audiences amongst the TV public in recent years should not be taken to suggest that they will continue at this high level indefinitely. The amount of listening to sound radio, at least in recent years, has shown variations that would appear to bear little relationship to any changes in the intrinsic merit of the programme material. It is almost as though listening were a social "fashion"—at times it is the popular thing to do, it forms a major topic of conversation and so on. At other times it may be ousted from such a high place in the public eye by any of a number of other conflicting activities. No doubt the amount of viewing over the years will be subject to just such apparently arbitrary variations.

The figures in the last three rows of this Table 12 are comparatively low. They reflect the fact that TV audiences in the second and third quarters of 1955 were considerably below normal, probably because of the exceptionally fine summer. The quarterly figures are given in Table 13. They exhibit a fairly regular seasonal cycle—highest in the January/March quarter, somewhat lower in October/December, and distinctly lower in the two summer quarters.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 13

Quarterly Average Audiences amongst the Adult Television Public

	Average Daytime TV Audience %	*Average Evening TV Audience* %
1st Quarter, 1953	10·2	41·6
2nd ,, ,, ...	9·8	35·9
3rd ,, ,, ...	9·6	36·9
4th ,, ,, ...	9·9	41·7
1st ,, 1954	9·9	45·0
2nd ,, ,, ...	10·2	38·7
3rd ,, ,, ...	8·7	36·9
4th ,, ,, ...	9·7	41·6
1st ,, 1955	10·0	43·8
2nd ,, ,, ...	7·6	36·1
3rd ,, ,, ...	6·4	30·0
4th ,, ,, ...	8·4	36·6

The maximum audiences are generally found in February, when the evening average is about 45 per cent. of the television public and the minimum at the end of July (about 30 per cent. of the Television public). In the field of sound radio a precisely similar seasonal pattern has been observed for many years.

The sheer magnitude of these figures should be noted. During the course of a whole evening's transmission it is usual for between 60 per cent. and 80 per cent. of owner-viewers to see something. The normal "rate of consumption" per viewer is two evening broadcasts out of every five: which would be the same as saying that the average viewer watches, every week, nearly three evening transmissions from beginning to end. It should be remembered, too, that 40 per cent. is the average, not the maximum, proportion of viewers who see the evening programmes. Thus, in one typical week in which there were sixty evening television broadcasts, at the one extreme there were nine which were seen by less than 20 per cent. of the television public—most of them were late in the evening—and at the other, ten which were seen by over 60 per cent., including two with audiences which included three viewers out of every four.

A comparison of viewing and listening. Nothing more clearly emphasizes the "pulling power" of television than a comparison of the time which is spent in viewing with that which is spent in listening. Table 14 gives a picture of the position on a typical evening in the spring of 1954. It compares the behaviour of the television public at that time with that of the sound public (i.e. people with sound, but not television, receivers).

TABLE 14

Average Evening Audience—Spring, 1954

	Amongst Adult TV Public %	*Amongst Adult Sound Public* %
During the period when TV is *not* on the air:		
Listening to sound radio . .	16·6	27·4
During the period when TV *is* on the air:		
Listening to sound radio . .	3·8	18·6
Viewing Television . . .	42·6	1·0
	46·4	19·6

This table brings out several important facts:

(i) During the hours when sound and television were competing with one another the proportion of the television public which was viewing (42·6 per cent.) was more than twice as great as the proportion of the sound public which was listening.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(ii) The television public did not cease altogether to listen at times when they could view, but they listened very little—only about one-fifth as much as the sound public did (3·8 per cent. as against 18·6 per cent.).

(iii) Even at times when television was *not* on the air, the proportion of the television public listening (16·6 per cent). was much less than the proportion of the sound public doing so (27·4 per cent.).

Do these differences (between the listening of the television public and that of the sound public) represent the extent to which acquiring a television set reduces listening? The answer depends on whether the television public, before they became viewers, listened as much as the sound public is shown to do in Table 14. Of this there is no proof, but it seems highly likely that in fact the television public listened even more; for their purchase of a television set suggests that they attach high value to home entertainment and people who do so are *prima facie* likely to be big consumers of sound broadcasting. If this supposition is correct then the difference shown in Table 14 may actually *understate* the extent to which viewers reduce their listening.

Tables 11 to 14 dealt with the television public *as a whole*. The next point to be examined is the extent to which, within the television public, there are significant variations in viewing behaviour as between different types of viewer.

Viewing and the length of TV ownership. The relationship between the amount of viewing (and listening) and the length of time the viewer has had a set is clearly of paramount importance. It has never so far proved possible to record the viewing behaviour of the same sample of viewers over a number of years and so to arrive at firm conclusions about the trend of viewing "over time". The most that it has been possible to do is to collect and compare information about the viewing "at a given moment" of viewers who have had their sets for varying numbers of years.

The most recent information of this sort dates from January, 1954, but since it confirmed findings in 1948 and 1950 there is little reason to suppose that the present situation is materially different. In weeks 3–5, 1954, all Survey of Listening and Viewing informants who had television were asked to state when they first became viewers. An analysis of this material is given below.

TABLE 15

Amount of Viewing and Length of TV Ownership—Weeks 3–5, 1954

	Average Evening Viewing and Listening Audiences amongst Adults who first became Viewers in:				
	1953 %	1952 %	1951 %	1950 %	1949 *or earlier* %
Viewing	47·1	45·0	43·6	44·1	40·4
Listening (when TV on the air)	4·6	5·0	5·2	5·4	7·6
	51·7	50·0	48·8	49·5	48·0
Listening (when TV not on the air)	14·2	14·5	15·7	15·5	19·3

It is quite apparent from this table that the amounts of viewing and listening by viewers are associated with the length of time they have had television sets. The longer the ownership of television, the lower the level of viewing and, by contrast, the *greater* the level of listening, not merely during times when television programmes were available but also during those evening hours when the television transmitters were silent.

This evidence is certainly consistent with the hypothesis that, with the passing of years, viewers tend to reduce the amount of their viewing, though it should be noted that the extent of the decrease in viewing (if such it is) is not very great. The evidence is also consistent with the hypothesis that, parallel with a diminution in viewing as television sets get older, there is a reversion to listening—though the amount of listening during television hours, even by those who have had their sets five or more years, is still far below that of people who do not possess television sets. (These statements are made cautiously because in the absence of evidence collected "over time" it cannot be said with certainty that, for example, veteran viewers have *reduced* their viewing. All that is

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

known is that they were viewing less than new viewers when the data were collected.) Finally, the evidence indicates, as earlier studies had done, that the more the viewer views the less he listens, not merely during the hours when television programmes are available to him but also at other times.

It may be of interest to note at this point that there was nothing to suggest that these relationships—of listening and viewing to length of television ownership—affected one type of programme more than another. The audience of different kinds of television broadcasts all seemed to show the same tendency and, in sound, listening to the Home Service and the Light Programme seemed to be affected to a relatively equal extent.

The most that the three special television inquiries in 1948, 1950 and 1954 have disclosed in this respect is that persons who have been viewers for five or more years viewed less than did those whose purchase of a television set was more recent. A most interesting point for study would be to discover the viewing behaviour of those whose experience of television dates back even longer—say for ten years. Is there a sharper fall in the amount of viewing, or indeed a levelling out of the curve, beyond the five-year mark? Nothing is known of this—such veteran viewers being comparatively rare and hence difficult to sample. But without such information, the question "Whither television?" remains unanswerable. A reasonable guess might be that, since five years is, after all, quite a long time, the decline in the amount of viewing after that time is unlikely to be accelerated, and that viewing may well "settle down" at a level not much below the still very considerable rate reached at the end of five years. As was said earlier, there is no reason to expect stability in the amount of viewing even amongst the newcomers to TV and my guess would be that, as viewing becomes a normal and commonplace ingredient of life, the level of viewing will probably not much exceed say 25 per cent. or 30 per cent. instead of the present 40 per cent.

Viewing levels in B.B.C. Regions. Table 16 compares the levels of evening viewing in the seven B.B.C. Regions.

TABLE 16

Average Evening Audiences amongst the Television Public, by B.B.C. Region, during Year ended 31/3/55

	Average Evening TV Audience %
London Region	39·1
Midland ,,	40·2
North ,,	40·3
Scotland	43·1
Wales	42·9
West Region	38·7
Northern Ireland	39·4

It cannot be said that there is much difference in the extent to which viewers in different Regions use their television sets in the evening. The generalisation that "viewers see roughly two evenings' television programmes out of five" is broadly true everywhere.

Viewing and social class. There appears to be little difference between the extent to which television is viewed in the three broad social classes—the Upper Middle, the Lower Middle and the Working Class. The latest analysis on these lines related to Week 9, 1955, the results being given in Table 17.

TABLE 17

Average Evening Audiences amongst the Television Public by Social Class

	Average Evening TV Audience %
Upper Middle Class	42·5
Lower Middle Class	43·9
Working Class	43·9

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Social class, however is a notoriously multi-dimensional concept. It is, therefore, worth while seeing whether viewing is related to some of the components of social class. Some data are available about three such components, viz., occupational level, income of head of household and educational level. These data come from the 1954 inquiry in which viewers were asked about how many evenings a week they "usually viewed". Table 18 summarizes their answers.

TABLE 18

Frequency of Viewing of Adult Television Public

	Usually View	
	Up to 4 Evenings a Week %	*5 or More Evenings a Week* %
Occupational levels:		
Managerial and professional	43	57
Highly skilled	30	70
Moderately skilled	25	75
Semi- and un-skilled	21	79
Income levels:		
Over £15 per week	38	62
£10–£15 per week	25	75
£5–£10 per week	26	74
Less than £5 per week	22	78
Educational levels:		
Left school at over 18	45	55
,, ,, ,, 16–18	39	61
,, ,, ,, under 16	24	76

These figures need to be treated with some caution, resting as they do upon people's statements about the regularity of their behaviour. It appears to be a common experience of surveys that people who perform an action frequently are liable to overstate, and people who do it rarely to understate, its frequency. The present case provided evidence of this, for it was possible to compare people's statements about the numbers of evenings a week on which they usually viewed with the extent to which they had, in fact. viewed the previous evening. All those who claimed to view "every evening" would be expected to have viewed the previous evening; in fact fewer than 90 per cent. had done so. But of those who claimed to view not more than two evenings a week, 33 per cent. had viewed on the previous evening. Tendencies of a similar kind have been found in surveys of other fields, such as cinema-going and cigarette smoking. And even were these statements of frequency reliable, such figures do not necessarily reveal the *quantities* of viewing, for they take no account of how much of each evening's transmission is seen.

TABLE 19

Average Evening Audiences amongst the Television Public by Income and Education

	Average Evening TV Audience %
Income group:	
£20 per week or more	39
£13–£20 per week	44
£8–£13 per week	48
£4/10/– to £8 per week	49
Less than £4/10/0 per week	42
Educational background:	
University	44
Secondary	49
Elementary	51

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

More precise information on this point is obtainable from the 1950 inquiry when data were collected which enabled evening average audiences to be computed for each of five income and three educational groups. (See Table 19).

Reviewing the evidence as a whole there seems to be little doubt that the "lower" the viewer's status in terms of either occupation, income or education, the more he is likely to view. Nevertheless, the amount of viewing by those at the top of these ladders is so great that the difference between their viewing and that of viewers at the lower levels is not of great practical importance. Moreover, as the figures in Table 17 have shown, when a comprehensive classification of social status is employed the differences in the quantity very nearly disappear altogether.

Viewing and age. The 1950 television inquiry arrived at estimates, given in Table 20, of the average evening audiences amongst viewers in different age groups.

TABLE 20

Average Evening Audiences amongst the Television Public by Age

	Average Evening TV Audience %
Viewers aged:	
5–7 years	8
8–11	24
12–14	36
15	47
16–19	43
20–24	34
25–34	47
35–44	54
45–54	54
55–64	54
65 or more	51

The 1954 inquiry gave estimates of the numbers of evenings a week on which television was usually watched by adults in various age-groups that were broadly consistent with these figures. It therefore seems legitimate to regard the 1950 figures as giving a general picture which still holds to-day. They show evening viewing to vary considerably from age-group to age-group among children and young adults—increasing as they get older and then diminishing in the span from 16 to 24 years (when notoriously much leisure time is spent out of the home). Among adults over 35, however, viewing varies little with age.

I.T.A. Viewing. The impact of an alternative service on the amount of viewing the television public indulge in does not seem to have been very great. The mean viewing audience for evening programmes in the fourth quarter of 1955 amongst those with sets capable of receiving both B.B.C. and I.T.A. programmes was 36·9 per cent. Amongst the rest of the TV public who could not receive the alternative programmes it was a little larger, 38·1 per cent. In the United States, where most viewers have a choice of programmes—and some have a choice of as many as six or more—the average evening audience is also estimated to be of about this size. This may well imply that viewers are already devoting to television as much time as they are prepared to give between 7 p.m. and 11 p.m.

The Survey of Listening and Viewing estimate of the way in which viewers with a choice of programmes split their time between them is that, in the fourth quarter, 1955, the ratio of the average B.B.C. evening audience to the average I.T.A. evening audience was 54 : 46.* This conceals a steady trend throughout the quarter for the ratio to become gradually more favourable to the I.T.A. In the early weeks of the new service it was found that the B.B.C. audience amongst viewers who could receive both normally exceeded that of the I.T.A., but by the end of the year the reverse was true, the ratio of the average evening audiences in the last few weeks of the year being estimated at B.B.C. 49 : I.T.A. 51.

* It must be recalled that Survey of Listening and Viewing estimate refers to adults from TV homes in the London Region and for this reason alone cannot be expected to agree precisely with other estimates that have been made of the proportions of TV sets tuned to the B.B.C. and I.T.A.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

GUEST VIEWING

Viewing television is by no means confined to people who possess television sets. The 1954 inquiry showed that, of people *without* television sets, nearly one in three claimed to view evening television at least occasionally. (The average frequency of such viewing was no more than one night per week.) At that time the television public amounted to about one-quarter of the total adult population, and hence in the course of a typical week it would seem that television was being seen at some time or another by half the entire population (a quarter as owner-viewers and a quarter as their guests.)

Table 21 shows how, in the spring of 1954, the picture differed from income group to income group.

TABLE 21

Guest Viewing by Income Group—Spring, 1954

Income of Head of Household	*Proportion With TV* %	*Proportion Without TV: but who occasionally View as Guests* %	*Proportion Without TV: and who Never View* %	*Total* %	*Ratio of Guest-viewers to Owner-viewers* Guests : Owners
Over £25 per week . .	58	29	13	100	49 : 100
£15–£25 ,, ,, . .	34	28	38	100	82 : 100
£10–£15 ,, ,, . .	32	21	47	100	66 : 100
£5–£10 ,, ,, . .	22	26	52	100	122 : 100
Under £5 ,, ,, . .	13	22	65	100	171 : 100
All income groups . .	23	25	52	100	104 : 100

As might be expected, the higher the income group the smaller the proportion who never viewed either as owners or as guests. (In the top income group only 13 per cent. never viewed, whereas in the under £5 per week group the proportion who never viewed was 65 per cent.) On the other hand, the ratio of guest-viewers to owner-viewers tended to increase with each step down the income scale. In the highest income group there were only 49 guest-viewers to every 100 owner-viewers, and in the lowest, 171 guest-viewers to every 100 owner-viewers. The explanation is simple—the more common television sets are, the fewer are the people who are dependent for viewing upon being invited into the home of an owner-viewer friend. (The same sort of difference occurs between B.B.C. Regions with high and low densities of television. In the Regions where television is common, the ratio of guest-viewers to owner-viewers is low, and in Regions where television sets are comparatively rare there is a higher ratio of guest-viewers to owner-viewers.)

The trend shown by Table 21 can be used to suggest the size of the guest-viewing public at the end of 1955. The television public then amounted to about 40 per cent. of the total. With such a density of television the ratio of guest-to-owner viewers might be expected, on the basis of the trend in the income group figures, to be about 60 : 100. If this procedure of interpolating in the income distribution trend and applying the result to the total population has any validity, then it would seem that by the end of 1955 some two-thirds of the total adult population were at least occasional viewers (40 per cent. as owner-viewers and 25 per cent. as their guests).

Although the number of people who occasionally see television as guests is very large indeed, the frequency with which they view is so much less than that of owner-viewers that in the normal way they represent only small proportions of broadcasts' audiences. For example, in the spring of 1955, when the average of evening audiences was 14·9 per cent. of the adult population of the United Kingdom, 14·1 per cent. were owner-viewers and 0·8 per cent. were guest-viewers. In other words, out of every 100 people viewing, 94 were owner-viewers and only 6 were guest-viewers. Naturally enough it is on the "special occasion" that the owner-viewer is most likely to throw his home open to his friends. For example, out of every 100 people viewing television's Annual Party (Christmas, 1954), 78 were owner-viewers and 22 were guest-viewers; out of every 100 viewing the Cup Final, 72 were owner-viewers and 28 were guests.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ACKNOWLEDGEMENT

In conclusion I am pleased to acknowledge my indebtedness to my colleagues in the Audience Research Department, most of whom have contributed greatly to the research here reported, and several of whom have assisted in the preparation of the paper.

SUMMARY

Both in the United Kingdom and in many other countries in the world, television is in a state of rapid expansion. Already by the end of 1955, the television public in the United Kingdom included 40 per cent. of all adults. In the United States and Canada the proportions were even higher. Television has been acquired more rapidly by the well-to-do; but in the past year or two the rate of growth of the television public in the lower income groups has increased, so that quite soon it may be expected that the television public will differ very little in income structure from the population as a whole. In terms of size and age composition, television households still differ from others, being considerably larger and containing considerably more 5–15-year-olds. The amount of time spent in viewing by the television public is substantial, the average of evening audiences being as much as 40 per cent. (It may be noted that in the United States and in Canada the average evening audiences are also of about this size.) There is evidence that the amount of viewing decreases as the novelty of television wears off, but even amongst persons who have been viewers for five years or more there does not seem to be a fall of more than about one-sixth. Viewing is by no means confined to members of the television public. Early in 1954 about one-quarter from amongst the three-quarters of the population who had *not* got television also viewed occasionally; though for them it was a comparatively rare event and they normally constituted only a small minority of the audience for television.

References

General Post Office, Monthly Statements of Wireless Receiving Licences Current (January 1947 onwards).
Hulton Readership Surveys, London, The Hulton Press (1949 onwards).
Television. A World Survey. Paris, U.N.E.S.C.O., 1953.
Television. A World Survey. Supplement 1955. Paris, U.N.E.S.C.O., 1955.

Bibliography

A fairly comprehensive bibliography of television research containing nearly two hundred items has recently been published, (1) below, and there are two others which contain additional material, (2) and (3) below. Most of the work to which they refer is American.

(1) ADAMS, J. STACY (1955), *A Selected Bibliography of Research in Television.* Chapel Hill, University of North Carolina.
(2) PAULU, BURTON (1952), *Radio-television Bibliography.* Urbana, Illinois, National Association of Educational Broadcasters.
(3) —— (1954), *Bibliography for Listeners and Viewers*, Los Angeles, National Association for Better Radio and Television.
(4) BELSON, W. A. (1955), "Learning and Attitude Changes Resulting from Viewing a Television Series 'Bon Voyage' ", *Brit. J. Educ. Psychol.*, **26**, 31–38.
(5) DOWNHAM, J. & CAUTER, T. (1954), "The Communication of Ideas". London, The Readers Digest Association.
(6) SILVEY, R. J. E. (1955), "Television Tips the Scale". *The Times Radio and Television Supplement*, 19th August, 1955.

DISCUSSION ON MR. EMMETT'S PAPER

Mr. A. STUART: Mr. Emmett has presented us with a very clear summary of what may be called census-type data for the adult television audience. A good deal of the material is not of current interest, in the sense that, as the paper emphasizes, the social penetration of television has now for the first time reached the point where its audience begins to approximate the general population in its demographic characteristics. But the paper is not the less valuable for this.

I imagine that, despite the B.B.C.'s earlier publications, quite a few Fellows remote from the television field will have been surprised to learn that the television audience has spread steadily downward, and not upward, through the social scale: the idea dies hard, at least among my immediate acquaintance, that television aerials sprout only on rent-controlled roofs. (In fact, a popular political joke when the Purchase Tax is increased is that the Chancellor of the Exchequer

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is grinding the faces of the poor into their TV screens.) Mr. Emmett shows that the truth behind this fallacy is that, at a given income level, the less-educated tend to acquire television more rapidly. The fallacy is an interesting example for the statistician of confusion between total and partial correlation, a confusion probably fostered by the physical preponderance of the less-educated.

Mr. Emmett has not discussed automatic recording methods for viewer research, which the B.B.C., unlike some commercial organizations, does not use. I cannot agree that these devices have nothing to contribute to viewer-research. Even if only as checks on memory reliability in personal interviews, they are of value; but I believe their major contribution to viewer research will be in the continuous nature of the records they provide. Used on an adequate scale, in conjunction with the interview method, they could considerably increase the precision of estimated changes in the television audience. For example, the paper's conclusions on "veteran viewing", below Table 15, might have been made more precise by such methods.

A most interesting feature of the development of television in this country has been the serious concern of responsible persons and bodies at the medium's potent possibilities. Commercial broadcasting has increased this concern, but it was sufficiently well developed before. Whatever its origins, such concern must of itself be held to justify a deepening of the B.B.C.'s studies in this field. As far as I know, the B.B.C. is not bound to make its audience research cover the whole wide field of broadcasting, and I should like to see its routine research (as distinct from the special studies to which Mr. Emmett has referred) incorporating more intensive study of particular research sectors.

I have pleasure in proposing the vote of thanks to Mr. Emmett for his paper.

Mr. H. F. Lydall: Mr. Emmett has given us a clear and interesting account of the growth of the television audience in the United Kingdom, of the nature of the television public, and of the number of hours spent in viewing. The figures which he presents are very remarkable. Is it not a humbling thought, for example, that at the end of 1955 there were 15 million adults who were spending an average of 11 hours a week watching television? In four years' time, according to Mr. Emmett's prediction, there will be nearly twice as many people with television sets and presumably they also will be spending the same average amount of time on it. It is clear that television is an important potential source for good or evil in our society; and it is not surprising, I think, that people should be concerned about it. Nor is it illogical, as has sometimes been suggested, for people who have accepted the inevitability of advertising in other media to question the desirability of its extension into this new field.

I must confess that the questions that interest me most about television are those which Mr. Emmett has deliberately left on one side. It is, of course, important to know how many people have television sets, what income and social classes they belong to and how much time they spend on this form of recreation. What seems to me to be far more interesting and important is to know what effects television has on the people who view. What kinds of programmes do they prefer? What have they learnt from seeing them? What effects have these experiences had on their tastes and on the cultural pattern of their lives? I am very glad to hear that research on these topics is now in progress and I hope that the results will soon be made public.

I have one or two detailed comments on Mr. Emmett's paper. First, if we consider the graph in Fig. 4, showing the growth of the total television audience, we must be struck by its similarity to a typical growth curve of the logistic type. If the curve continues to conform to a logistic shape it will pass through a point of maximum slope and begin to increase more slowly thereafter. There seem to be good *a priori* reasons to expect such a development, although we cannot confidently predict the upper level to which the curve will be asymptotic. Will television follow radio so that eventually it will be installed in virtually every home? Or is there a hard core of the population who will, for one reason or another, never acquire a set?

Secondly, on p. 297 Mr. Emmett shows that the proportion of television owners amongst one-person households is much lower than amongst other households. Many of these households are single old-age pensioners and they are precisely the sort of people whom one would expect to want to own a television set. If the price of television sets comes down enough, or the incomes of old-age pensioners go up, we may surely expect these one-person households to acquire television sets on the same scale as the rest of the population.

Thirdly, on p. 301 Mr. Emmett discusses the evidence that, as the years pass, viewers tend to reduce the amount of time spent on viewing. Although this conclusion is not surprising, the evidence is in fact susceptible of another interpretation. The people who acquired television sets in 1949 or earlier were, as Mr. Emmett has shown, more commonly of the upper and middle income groups than those who bought their sets in more recent years. People in the upper income

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

groups have more freedom of choice in their leisure time occupations and may, for this reason alone, if not for others, spend less time at their television sets than those who are less well off. This tendency is confirmed by the figures in Table 18.

I have much pleasure in seconding the vote of thanks.

The vote of thanks was put to the meeting and carried unanimously.

Professor P. E. VERNON: I am sure you will forgive me if I speak more as a psychologist than as a statistician. I think it is appropriate that I should do so since a good deal of the interest of this valuable paper is clearly psychological and sociological. It tells us what kinds of people own sets and view, and throws light on their habits and tastes. Not only does Audience Research provide the B.B.C. with the routine data it needs, but also it is very much alive to the desirability of fundamental research, for example, into techniques of collecting information from the public, and into the short- and long-term effects of various kinds of programmes; and Mr. Emmett and his colleagues have, with the limitations of staff and resources, made a considerable contribution along these lines.

We all know that interviewers, particularly those not highly trained or well paid, introduce all kinds of bias into their samples and into the answers of their informants, and perhaps Mr. Emmett might have spent more time on telling us of the checks Audience Research carries out.

Recently there was a very enterprising experiment in which large numbers of houses were visited within 60 minutes after programmes had been broadcast, and the numbers thus found to have heard various programmes checked quite reasonably with the normal interview returns collected on the following day. Although that experiment related to sound broadcasting, it does seem important for TV also in view of the recent attacks of various organisations on the B.B.C.'s published figures.

If there is a point in the paper that I might perhaps disagree with it is that Mr. Emmett possibly underestimates the effect of social class differences. He says that such differences are not of great practical importance. It is true that set ownership is little affected nowadays by class, but his figures do appear to show considerable differences in viewing habits among the more cultured classes. Perhaps I might ask why in Tables 7 and 8 he compared TV social class distributions with sound distributions rather than with the general population.

Finally, I was a little puzzled at his prediction that the ultimate level of viewing may be of the order of 25 to 30 per cent. It seems to me that TV becomes an established habit, very much like filling in the pools, or the Sunday joint, and that therefore the figure of 40 per cent. which is indicated in Table 15 is more likely to remain close to the minimum figure. I should be very interested to know on what grounds Mr. Emmett thinks there will be a considerable decline below that figure.

Mr. E. J. DAVIS: I belong to Television Audience Measurement; I am on the other side of the fence. I do not want to start arguments on the figures, for as far as our own data go, we agree. Mr. Emmett has been engaged in measuring the television audience for much longer than we have, but the results obtained in our six or eight months of life would in general confirm his.

The B.B.C. have been working on quota samples since 1947. I imagine that over that time the basis on which the quotas were constructed must have changed. In 1947 there was not a great deal of information available, and it was not until 1952, I think, that we had any census data. I should like to know to what extent the B.B.C. have had to modify their quotas over the nine years, and whether this is likely to have any effect on the figures.

I was rather pleased that Mr. Emmett put in the seasonal picture, because some people have tended to ignore it, and have attributed the increase in viewing since September to the opening of the I.T.A. stations.

Mr. Emmett has provided some very interesting information on the distribution of television sets in the London area, but this would be even more interesting if a map of the six districts could be included.

Now to the crux of the matter. Mr. Emmett says that towards the end of 1955, 51 per cent. of viewers able to get both programmes were watching I.T.A. and 49 per cent. were watching B.B.C. programmes. I am not going to quote figures but I would suggest that our results were nearer 60 per cent. and 40 per cent. That has led to a tremendous amount of controversy and I think, having looked at this from a lot of angles, that basically it is the old problem of two different organisations putting figures against two different definitions. In our work we use random samples, we visit households, and we assess for ourselves the quality of reception. Unless reception is satisfactory the household is not included in our results as receiving both programmes.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

There are then, outside the households we classify as being in the I.T.A. audience, numbers of households in the fringe areas, or other areas of doubtful reception, which can probably receive programmes on perhaps one or two evenings a week. They can get the programmes, but not sufficiently well for us to include them in the I.T.A. audience on our definitions. Consequently we would tend to show a higher percentage of people watching I.T.A. at any time than would the B.B.C. research figures, since the B.B.C. work covers a wider area, and is based on a less rigorous definition of I.T.A. reception.

I think rather too much is being made of the differences in the figures. We are probably both right in our results, but are talking about slightly different things.

Mr. A. GABOR: This paper contains a wealth of useful information which will help future research.

I myself am engaged on an analysis of certain data made accessible to me by the courtesy of the British Broadcasting Corporation. As I hope to present the full results at a later date, I will restrict my observations to two of the points made by Mr. Emmett.

The measures of average audiences given in Tables 12 and 13 are appreciably lower than those I have computed by using a somewhat different definition. The following figures compare with the appropriate entries of Table 13:

	Average Evening Adult TV Audience (%)
1st Quarter, 1953 . . .	60·2
2nd ,, ,, . . .	53·4
3rd ,, ,, . . .	48·6
4th ,, ,, . . .	56·4
1st ,, 1954 . . .	59·0
2nd ,, ,, . . .	49·3
3rd ,, ,, . . .	45·9
4th ,, ,, . . .	54·5
1st ,, 1955 . . .	56·8
2nd ,, ,, . . .	47·2

For the period covered by the above figures the straight average of Mr. Emmett's entries is 39·8 per cent., whereas mine give 53·1 per cent. The discrepancies may be due to the following differences in the definitions used:

(i) Mr. Emmett's figures refer to adult members of families with a television set, whereas mine are based on returns which include guest viewers, (except on certain evenings where the proportion of guests was particularly high, and the returns supplied additional estimates of net numbers). However, Table 14 implies that this is unlikely to account for more than a small fraction of the discrepancies.

(ii) Mr. Emmett's figures cover the whole of the evening, including "sound only" programmes, which were excluded from my analysis. I further excluded all repeats, as also items for which the information was in some respect incomplete.

(iii) Mr. Emmett's figures are straight averages, whereas mine are weighted by the duration of each programme. There is a strong indication of the presence of appreciable correlation between duration and audience numbers, and this may well be the major factor responsible for the discrepancies.

I would conclude that the actual proportions of viewers of programmes not presented before were on the average about one-third higher than is implied by Mr. Emmett's data. Similar considerations apply to the time people appear to spend viewing, though it should be remembered that the returns on which both Mr. Emmett's figures and mine are based make no allowance for those persons who viewed part of the programme only.

The slight downward trend to which Mr. Emmett refers appears also in my analysis. The crude trend is difficult to interpret, because it is the outcome of at least two conflicting influences. The changing structure of the TV public grouped according to duration of ownership of a TV-set implies a *downward* trend (with the qualification mentioned by Mr. Lydall), but since the average income and level of education of the TV public have fallen at the same time, the implication of Table 19 is that an *upward* influence has also been at work. With the information available in respect of past periods there is little possibility of clarifying this important issue to any great extent, but I hope that future enquiries will provide the relevant data.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Mr. WOOD: On p. 288 there is a reference to the Post Office figures of licences. The statements there are not quite up to date. The figures we in the Post Office produce now, both monthly and quarterly, come from the same source, the Head Postmasters' Records, and are therefore consistent.

Dr. BARNA: I very much enjoyed listening to Mr. Emmett's paper. There is one deficiency, however, which as a statistician I think I ought to mention: nowhere are margins of error given for the estimates, and that applies also to the contributions of the two speakers in the discussion who gave alternative estimates.

In more than one place the difference between social classes or income classes looks small, and Mr. Emmett interprets these differences as being practically unimportant, or practically important. It is not at all clear whether these are subjective judgments or based on some back-room calculation of sampling errors. I think that at least the straightforward sampling errors should be given, especially as some of the numbers involved when the data are cross-classified are relatively small.

It is clear that there are other kinds of errors also. I can think of the error arising in quota sampling, when it is not certain that the person interviewed is in social class 1, 2 or 3, and of the errors involved in interpreting the concepts. It might be difficult to put a quantity on all these different errors, but I should very much like to have some quantitative interpretation of these factors.

Mr. J. A. BOUND: I was much struck by Fig. 3, showing the density of television in various parts of the country. Mr. Emmett suggests that part of the influence may be due to variations in field strength, but if the position of stations is automatically in or close to areas of dense population this may be merely an effect rather than a cause. This still leaves the point that the density of television seems to be very much greater in areas with the highest populations. Is this due to some variation in the social structure in the various parts of the country or to some other factor?

Mr. J. I. MASON: The need for figures of the distribution of population by Post Office areas is becoming increasingly acute. I wonder whether some action could not be taken officially to provide estimates for each Post Office area centrally. Even if only approximations were provided, they would be enough for many researchers to work on, who at present have to make tedious "guestimates" in a most unsatisfactory way.

Mr. EMMETT subsequently replied in writing as follows:

I am glad that Mr. Stuart referred to the fallacious notion that television has been acquired most rapidly by the lower social classes. I have often had to correct this misunderstanding myself. His desire for further research in the field of effects of television, strongly pressed also by Mr. Lydall, is shared inside the B.B.C. and considerable strides have already been taken in this direction not only by the B.B.C.'s Audience Research Department but by others. Within a year some important results should be available. My speculation as to the future growth of television cannot, I think, be made more precise by a consideration of the logistic type curve suggested by Mr. Lydall. The recent credit squeeze has already, it seems, slowed down the rate of growth of the television public and thus distorted the curve. Future occurrences may have even more marked effects. The "hard core" of people who will never acquire a television set no doubt exists but in my opinion is likely to be very small. Any attempt to forecast the position more than four or five years ahead would seem to me quite futile. Professor Vernon mentioned the validation work we have carried out to satisfy ourselves about the adequacy of the methods we use in our routine Survey of Listening and Viewing. Generally speaking, we have been well satisfied with the results. No general publication has been envisaged but I should be pleased to discuss this matter with anyone interested in audience measurement methods. His second point refers to Tables 7 and 8 in which comparisons are made between the "TV public" and the "sound public". The "sound public" rather than the general population was used because in this way the contrast is most vivid. Virtually all who have no TV set have a radio set, so that the general population figures would be weighted means of the TV-and sound-public figures, thus lying between the two. Professor Vernon's final comment refers to my guess (I prefer that word to prediction) at the ultimate level of viewing in the future. He may well be correct in thinking that the 25–30 per cent. I gave is too low. I am confident it will fall below the 40 per cent. current until 1955 (*a*) because the proportion of "new viewers" (who view most heavily) must gradually fall and (*b*) because the level of interest in TV cannot, I think, be maintained indefinitely. The time must surely come for instance when newspapers no longer carry long daily TV reviews and gossip columns. When it does, the habit of spending a great deal of time viewing may sustain the level for a while but not indefinitely.

Mr. Davis touched on the controversy that has raged about differences in results achieved by different organizations in the TV audience measurement field. I entirely agree with his concluding

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sentences. All of us in this field will no doubt gain by discussing the differences in our results but it will probably be found that they are due far more to differences in what is being measured than in inadequacies of the methods. He asks what changes have been made in the quotas we set. They are of course revised periodically and it is not necessary to wait for Census data to do this. Estimated sex and age distributions are available at much more frequent intervals and so are figures of the numbers of gainfully occupied persons. The only other quota control employed is social class, and here we have generally found it far better to estimate the distribution from random sample surveys than to attempt an estimate based on Census data. He also asked for a map of the six zones into which we divide the London Region for the purposes of the Survey of Listening and Viewing. I regret that I have not been able to prepare one. The zones are designed solely for internal administration and I might have been wiser to omit Table 4 entirely. However, having included it, here is a short description of them: Zones I and II are slightly more extensive than Greater London, Zone I being north of the Thames and Zone II south of the Thames. Zone III extends in a band around the north of Greater London reaching out to Chelmsford, Hitchin and Oxford. Zone IV includes all Sussex and those parts of Surrey and Kent not included in Zone II. Zone V is another band running across East Anglia north of Zone III from Ipswich to Huntingdon and Zone VI is the rest of East Anglia.

Mr. Gabor notes the substantial difference between the arithmetic means of the audiences of all evening TV items (such as I gave in Table 12) and the means he derived by "weighting" audiences according to the duration of the programmes and excluding certain items. It was to be expected that such a procedure as he adopted would yield larger values than the ones I gave but I was interested to see that they turned out to be so very much larger. The extraordinarily high rate of "consumption" of new material (i.e. excluding "repeats" etc.) is clearly seen.

I am grateful to Mr. Wood for correcting my statement about the General Post Office returns. It is some time since I used the monthly figures and I had not realized that they and the quarterly returns were now "in step". Dr. Barna has put his finger on a manifest weakness in the results I presented, namely the lack of any standard errors. The Survey of Listening and Viewing from which most of my data were derived is, of course, based on quota samples and I would hesitate to guess at the sampling errors involved. Even the special television enquiry in 1954 was so unorthodox in sampling design that valid estimates of error could not be calculated. He, and others, may perhaps care to make rough estimates, as I have done in the past, based on the sizes of the samples (which are given in the paper) but they cannot be very dependable. My judgment as to the significance of such differences as were quoted is indeed far more a subjective judgment than one based on such rough estimates; though there is one firm foundation for such judgments, namely the consistency of the results from successive surveys made by different methods and even by different organizations. The only other point I need to comment on is that "practical importance" (which, incidentally, I used only once) means practical importance to the B.B.C. Other users may of course place different emphasis on the figures. As Mr. Bound said, the reasons for variations in the density of TV from place to place form an interesting topic for speculation. I cannot see how the influence of variations in field strength can be affected by density of population (persons per square mile) and in fact it is by no means always areas of high density of population that show the highest density of TV. Social class differences may well account for some of the variability but I think that a whole complex of factors would need to be examined before any definite conclusion could be reached. I am interested to learn from Mr. Mason that information about Head Post Office Districts would be valuable in other research as well as our own. I strongly endorse his plea for such information.

As a result of the ballot taken during the meeting, the candidates named below were elected Fellows of the Society:

Stanislaw Babik.
Keshab Chandra Banerjee.
Richard Stamp Benson.
Bernard Walter Douglas Bransden.
James William Cameron.
Richard Clowes.
Sylvia Dick.
Ana Maria Flores.
Armando Hagatong.
Hyman B. Kaitz.
Andrew Thomas Jamieson.
Anil Chandra Nag.
Ernest Arthur Powdrill.
Chiu Hock Quan.
William Patrick Quinlan.
Frederick David Robinson.
Harold Bertram Rose.
George Stanley Sykes.

Corporate Representative

Tudor Picton Miles, *representing* The Iron and Steel Board.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Some Theory of Index Numbers

By You Poh Seng

Introduction

Of the subjects covered by Statistics probably the least studied is that of index numbers. Any discussion of index numbers almost invariably includes the following topics: (i) An examination of the various aspects to be considered in the compilation of index numbers, such as the items or commodities for inclusion; the basic data in the form of prices, wages, etc.; the choice of base year; the weighting system; and the method of construction. Of these, all except the last can be discussed from a strictly practical standpoint. The method of index number construction is then developed to form the next topic. (ii) An enumeration (with explanation) of the many methods used in the construction of index numbers, (iii) a comparative appraisal of these methods, (iv) other aspects of construction, such as chain base *versus* fixed base, moving weights, link relatives, and other more or less specialized problems.

Topics (i) and (ii) have generally been fairly comprehensively and well treated, but (iii) and (iv) relatively neglected. The comparison of the methods of construction has been based on two properties: first, the practical property, such as ease of understanding, ease of computation and the like; second, the theoretical property. This latter is about the least developed branch of index numbers, for nearly all writers on the subject have hinged their discussion on the two or three "desirable" properties which were first put forward by Irving Fisher, namely, the Time Reversal Test, the Factor Reversal Test, and sometimes the Circularity Test which is, in fact, an extension of the Time Reversal Test. No progress has been made since Fisher's time in the search for other properties which an index number should possess. Even with regard to the standard tests, the general line of approach is as follows: only one index number (that using the simple geometric mean of relatives) obeys the time reversal test and the circularity test, and only one (the Fisher "Ideal" Index) obeys the factor reversal test; but since these are only desirable properties the non-possession of which is of no very great import to the usefulness of any index number, nothing further need be said about them. Though this may be true, it leaves the choice of method of construction very vague, and dependent almost entirely on the practical considerations of ease and convenience.

Statement of Aims

The aims of this thesis are: (*a*) to set out the results of an investigation into other properties of an index number; (*b*) to explain the advantages of the possession of such properties; (*c*) to discover under what conditions the various index numbers possess these properties, if at all; (*d*) to review the resultant relative position of the various index numbers; and (*e*) to consider the effect of the other specialized aspects of index number construction, such as chain base, linking, etc., on the new properties.

These properties can be briefly stated as follows:

(i) The maintenance of continuity of index numbers with changes in the base year;
(ii) The maintenance of continuity when one or more items (or commodities) are substituted for the equivalent number of old items;
(iii) The maintenance of continuity when one or more items are added to the original list of items;
(iv) The maintenance of continuity when one or more items are deleted from the original list of items;
(v) The maintenance of continuity with a change in the weighting system; and
(vi) The maintenance of continuity when two or more of the changes considered above take place simultaneously.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Hardly one index number compiled at the present time can claim to have retained all its original aspects of compilation such as the base year, the coverage, the weighting system, and sometimes even the method of construction (namely, the kind of average used). The simplest change is in the base year, but in most cases this is itself a signal for the introduction of other changes. Or it may be the other way round: the need for a change in the coverage, etc., generally brings with it a change in the base year.

One problem in using any one index number for comparability over time or for time series and other econometric analyses is to obtain a long enough series for the index numbers to be comparable throughout the required period. In each of the above situations, continuity can always be maintained by simply recalculating the series extending backwards in time. But there are many reasons why this is generally not possible. The labour involved constitutes the most important single reason, especially where the series is long or where there have been several changes, for example, in the base year of construction. In addition, the original data used in the compilation years ago may not be readily available, necessitating further wastage of time and resources in their re-assembly. Sometimes such data are not available at all.

These properties of continuity were touched upon when, in 1935, the United Kingdom Board of Trade, in using the geometric mean method of construction in the compilation of the Index Number of Wholesale Prices, claimed for this method the first four of the above properties.* These properties are stated without proof or discussion. Moreover, the implication that the simple geometric mean method of construction is the only method possessing any or all of these four properties will be seen in the course of the following discussion to be not quite true and probably the result of only a superficial examination of the matter.

It may be asked why the matter has not been studied thoroughly before this. The answer is that the need for maintaining continuity has not yet been sufficiently felt in practice. Most index numbers at present in use have not been in existence long, and important changes in base year, coverage, weighting, etc., have only begun to take place since the thirties. Users of index numbers have been chiefly concerned with the present as portrayed by such statistics, and where comparison is involved it is usually with the very recent past.

The need for continuity will undoubtedly be felt when any index number has been in existence long enough for use in time-series and similar econometric analyses. The major problem at the present moment in time-series analyses is the lack of sufficiently long series, and when the series is long enough, the question is bound to arise: how to make the series continuous with all the changes in the various aspects of construction. It would surely be wise either (*a*) to approach the matter directly and consider the best method of constructing an index number so that it not only presents a reasonable picture of the present, but can also fairly easily be made continuous; or (*b*) to approach it indirectly as follows: an index number is already constructed by such and such a method; can it yield a continuous series without too much labour in reassembling the necessary data and in recalculating, in view of all the changes that have taken place in the various aspects of its construction?

The commonest method of obtaining continuity is by a simple ratio process as follows: I is an index number compiled from say year 0 to year t. In year t some or all the aspects of compilation of the index number underwent some changes resulting in a new index I'. For year t both I and I' are available, but for years before t, i.e., for years $(t-q)$ $(q = 1, 2, \ldots, t)$, only I is available. The ratio process postulates proportionality of the following form:

$$\frac{I'_{t-q}}{I_{t-q}} = \frac{I'_t}{I_t}$$

so that for the years before t, I' can be obtained easily as

$$I'_{t-q} = \left(\frac{I'_t}{I_t}\right) \cdot I_{t-q}.$$

In this way the series of index numbers is made continuous from year 0 to beyond year t.

* *Board of Trade Journal Supplement* 24.i.35.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

This is probably the simplest method of maintaining continuity if the proportionality condition holds; it, in fact, obviates the necessity of collecting data for past years. In a further paper this method will be examined in detail in connection with the various changes in the method of compilation enumerated above.

In the present paper, except for base-year changes, the discussion will be concerned with obtaining continuity without the assumption of proportionality of the new index number to the old. It will be seen that a certain amount of past data will be necessary; the aim is to minimize this requirement.

Review of the Methods of Index Number Construction: Symbols and Notations

Twelve methods of index number construction will be discussed in this and in the following paper. There are several other methods, such as those involving cross-weights and the like, but it is believed that these twelve form more or less the basic system for a first study.

For the sake of convenience, only an index number for prices will be considered and all through, P will be used to denote price, and Q quantity. There will be N items or commodities, and any item will be denoted by a prefix, leaving the suffix to indicate year or time. Thus ${}_iP_t$, say, will mean the price of the ith commodity in year or time t. Again for convenience sake, all the summation and multiplication symbols, unless otherwise specified, will denote such operations over the range of commodities i ($i = 1, 2, \ldots, N$). a will be used throughout to indicate an arbitrary year.

The twelve methods of construction are further grouped under two families, and are shown in detail as follows (all index numbers will here be assumed to be based on year 0 = 100):

A. Index numbers involving some average of relatives:

1. Simple arithmetic mean:

$$I(\text{A.M.}) = \frac{100}{N} \sum_i \frac{{}_iP_t}{{}_iP_0}$$

2. Simple geometric mean:

$$I(\text{G.M.}) = \sqrt[N]{\prod_i \frac{{}_iP_t}{{}_iP_o}} \times 100$$

Weighted arithmetic mean, of three kinds, using as weights:

3. Base year values:

$$I(\text{W.A.b.}) = \frac{100}{\sum_i {}_iP_0 \cdot {}_iQ_0} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_0 \cdot {}_iQ_0$$

4. Current year values:

$$I(\text{W.A.c.}) = \frac{100}{\sum_i {}_iP_t \cdot {}_iQ_t} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_t \cdot {}_iQ_t$$

5. Arbitrary year values:

$$I(\text{W.A.a.}) = \frac{100}{\sum_i {}_iP_a \cdot {}_iQ_a} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_a \cdot {}_iQ_a$$

Weighted geometric mean, of three kinds, using as weights:

6. Base year values:

$$I(\text{W.G.b.}) = 100 \times \sum_i {}_iP_0 \cdot {}_iQ_0 \sqrt{\prod_i \left(\frac{{}_iP_t}{{}_iP_0}\right)^{{}_iP_0 \cdot {}_iQ_0}}$$

7. Current year values:

$$I(\text{W.G.c.}) = 100 \times \sum_i {}_iP_t \cdot {}_iQ_t \sqrt{\prod_i \left(\frac{{}_iP_t}{{}_iP_0}\right)^{{}_iP_t \cdot {}_iQ_t}}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

8. Arbitrary year values:

$$I(\text{W.G.a.}) = 100 \times \frac{\sum_i {}_iP_a \cdot {}_iQ_a}{\sqrt{\prod_i \left(\frac{{}_iP_t}{{}_iP_0}\right)^{{}_iP_a \cdot {}_iQ_a}}}$$

B. Index numbers involving ratio of aggregatives, having as 'basket of goods':

9. Base year quantities:

$$I(\text{Laspeyres}) = \frac{\sum_i {}_iP_t \cdot {}_iQ_0}{\sum_i {}_iP_0 \cdot {}_iQ_0} \times 100$$

10. Current year quantities:

$$I(\text{Paasche}) = \frac{\sum_i {}_iP_t \cdot {}_iQ_t}{\sum_i {}_iP_0 \cdot {}_iQ_t} \times 100$$

11. Arbitrary year quantities:

$$I(\text{Agg.a.}) = \frac{\sum_i {}_iP_t \cdot {}_iQ_a}{\sum_i {}_iP_0 \cdot {}_iQ_a} \times 100$$

12. Fisher's ideal Index:

$$I(\text{F.}) = \sqrt{I(\text{Laspeyres}) \times I(\text{Paasche})}.$$

SECTION ONE

Maintaining Continuity with Changes in One Aspect only of Index Number Compilation

A. *Maintaining Continuity with Changes in the Base Year*

The problem can be stated as follows: An index number, I_{t0}, has been computed on base year 0 = 100, that is, for every year 1, 2, . . . , t there exists an index number (0 = 100). The base year is now to be changed to year t, say. After this year all index numbers will be computed on the new base; but how can those before year t be converted to index numbers based on $t = 100$ without recalculating, or using original data?

For this purpose the best procedure is that of using ratios. The ratio of the index number for year $(t - q)$, $q = 1, 2, \ldots, t$, on base year 0, to that of year t on base 0, gives the index number of year $(t - q)$ on base t. Symbolically:

$$\frac{I_{(t-q)0}}{I_{t0}} = I_{(t-q)t}.$$

Only three of the twelve types of index numbers possess this property of ratio continuity, namely:

(a) The simple geometric mean index number, I(G.M.), where—

$$\frac{I_{(t-q)0}}{I_{t0}} \cdot 100 = \frac{\sqrt[N]{\prod_i \frac{{}_iP_{t-q}}{{}_iP_0}}}{\sqrt[N]{\prod_i \frac{{}_iP_t}{{}_iP_0}}} \cdot 100 = \sqrt[N]{\prod_i \frac{{}_iP_{t-q}}{{}_iP_t}} \cdot 100 = I_{(t-q)t}.$$

(b) The weighted geometric mean index number using arbitrary year values as weights I(W.G.a.), and

(c) The ratio of aggregatives index number using an arbitrary year "basket of goods", I(Agg.a.).

In (b) and (c) above the proof of the relationship, which is not given, is similar to that for (a).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

If the compilation of a new index number is being considered and continuity is an important factor, it might be suggested that either of the above three types be used. However, for index numbers already in existence, and not constructed by any of these methods, continuity may still

TABLE 1

Arithmetic Mean Method of Constructing Index Number: Showing the Ratio Continuity Conversion where the Base Year is changed from 0 *to* t. (*See text for explanation of the various cases here presented.*)

		Commodities									
		A	B	C	D	E	F	$I_{(t-q)0}$	I_{t0}	$\frac{I_{(t-q)0}}{I_{t0}}$	$I_{(t-q)t}$
	P_0	1/–	2/–	2/–	1/–	1/–	4/–	—	—	—	—
	P_{t-q}	1/6	2/6	3/6	1/–	2/6	6/–	—	—	—	—
	$\frac{P_{t-q}}{P_0}$	150	125	175	100	250	150	$158\frac{1}{3}$	—	—	—
Case 1	P_t	2/–	4/–	4/–	2/–	2/–	8/–	—	—	—	—
	$\frac{P_t}{P_0}$	200	200	200	200	200	200	—	200	$79\frac{1}{6}$	—
	$\frac{P_{t-q}}{P_t}$	75	$62\frac{1}{2}$	$87\frac{1}{2}$	50	125	75	—	—	—	$79\frac{1}{6}$
Case 2	P_t	1/–	2/6	4/–	1/–	1/3	4/–	—	—	—	—
	$\frac{P_t}{P_0}$	100	125	200	100	125	100	—	125	$126\frac{2}{3}$	—
	$\frac{P_{t-q}}{P_t}$	150	100	$87\frac{1}{2}$	100	200	150	—	—	—	$131\frac{1}{4}$
Case 3	P_t	2/–	1/6	4/–	–/6	2/6	8/–	—	—	—	—
	$\frac{P_t}{P_0}$	200	75	200	50	250	200	—	$162\frac{1}{2}$	$97\frac{17}{39}$	—
	$\frac{P_{t-q}}{P_t}$	75	$166\frac{2}{3}$	$87\frac{1}{2}$	200	100	75	—	—	—	$117\frac{13}{36}$
Case 4	P_t	1/6	1/6	3/–	2/–	2/–	5/–	—	—	—	—
	$\frac{P_t}{P_0}$	150	75	150	200	200	125	—	150	$105\frac{5}{9}$	—
	$\frac{P_{t-q}}{P_t}$	100	$166\frac{2}{3}$	$116\frac{2}{3}$	50	125	120	—	—	—	$113\frac{1}{18}$
Case 5	P_t	4/–	2/6	4/–	1/–	4/6	10/–	—	—	—	—
	$\frac{P_t}{P_0}$	400	125	200	100	450	250	—	$254\frac{1}{6}$	$62\frac{18}{81}$	—
	$\frac{P_{t-q}}{P_t}$	$37\frac{1}{2}$	100	$87\frac{1}{2}$	100	$55\frac{5}{9}$	60	—	—	—	$73\frac{23}{54}$
Case 6	P_t	4/–	1/6	4/–	–/6	4/6	5/–	—	—	—	—
	$\frac{P_t}{P_0}$	400	75	200	50	450	125	—	$216\frac{2}{3}$	$73\frac{1}{13}$	—
	$\frac{P_{t-q}}{P_t}$	$37\frac{1}{2}$	$166\frac{2}{3}$	$87\frac{1}{2}$	200	$55\frac{5}{9}$	120	—	—	—	$111\frac{11}{54}$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

be required. The question arises: under what conditions can this simple ratio continuity conversion be applicable or approximately applicable?

For the arithmetic mean method of construction, for example,

$$\frac{I_{(t-q)0}}{I_{t0}} \cdot 100 = \frac{\sum_i ({}_iP_{t-q}/{}_iP_0)}{\sum_i ({}_iP_t/{}_iP_0)} \cdot 100$$

which will be equal to

$$\frac{100}{N} \sum_i \frac{{}_iP_{t-q}}{{}_iP_t} \; (= I_{(t-q)t})$$

only if $P_0 \propto P_t$ for every commodity i.

To see the implication of the above result, Table 1 (shown on p. 316) has been drawn up to show some situations that are likely to arise in practice. The data are fictitious, and in each of the six cases shown, the P_0's and P_{t-q}'s are identical. The P_t's are the only factors that change, thus:

Case 1.—Where there is proportionality of the P_0's to the P_t's, the ratio continuity relationship holds exactly—i.e.

$$\frac{I_{(t-q)0}}{I_{t0}} = I_{I_{(-q)t}}.$$

Case 2.—Where there is no proportionality but the P_t's show consistent increases compared to the P_0's the equality is fairly close.

Case 3.—Where two of the P_t's show decrease and the remainder show consistent increase the equality is no longer approximate.

Case 4.—Where there is only one decrease, but the increases in the other commodities are not as consistent as in the previous case, the result is more or less similar to that case.

Case 5.—Where all the commodities show increases in year t compared to year 0 but the increases are very disproportionate, no equality or even approximate equality is noticeable.

Case 6.—Where there are increases and decreases, and all are very inconsistent, the inequality is intolerably large.*

In practice it is not to be expected that Case 1 will ever occur. Case 2 approximates closer to reality, although the other situations illustrated in the table and many others more or less similar are not uncommon.

Two observations can be drawn from this: (a) while it is true that approximate proportionality does not debar the use of the ratio continuity conversion, a fair departure from it does invalidate this procedure as it tends to give a misleading result. (b) The solution to the problem then is to be found in choosing the new base year t to observe, among other things, the following condition: that the prices in this new base year should be proportionate or fairly proportionate to those in the original base year 0.

The discussion so far concerns the arithmetic mean method of construction. Table 2 shows the conditions for the use of the simple ratio continuity conversion for the other methods of construction.

It is important to mention that in I(W.A.b.), I(W.G.b.) and I(Laspeyres), the change of the base year automatically involves the change of the weighting. I(W.A.c.), I(W.G.c.) and I(Paasche), being index numbers using current year weights, are not so affected. As for those index numbers

* It is to be noted that in each case from Case 2 to Case 6, the ratio gives a lower index than that obtained from $I_{(t-q)t}$. This result can be verified from the fact that the ratio $\frac{I_{(t-q)0}}{I_{t0}}$ is in reality equal to:

$$\left(\sum_i \frac{{}_iP_{t-q}}{{}_iP_t}\right) - \sum_i \sum_{\substack{j \neq i \\ k \neq i,j}} \left[\frac{{}_iP_{t-q} \cdot \frac{{}_jP_t}{{}_iP_t} \cdot {}_iP_0 \cdot {}_kP_0 \cdot \cdots}{{}_iP_t \cdot {}_jP_0 \cdot {}_kP_0 \cdot \cdots}\right],$$

which is thus less than $I_{(t-q)t}$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 2

Ratio Continuity Conversion for Changes in Base Year: Showing the Conditions under which this Conversion is Applicable for the Various Methods of Construction of Index Numbers. (In every case 0 *denotes the original Base Year, and* t *the new Base Year.)*

Method of construction	*Condition for Application of Conversion*
I(A.M.)	${}_iP_0 \propto {}_iP_t$
I(G.M.)	—
I(W.A.b.) . . .	${}_iQ_0 \propto {}_iQ_t$
I(W.A.c.) . . .	${}_iP_0 \propto {}_iP_t$
I(W.A.a.) . . .	${}_iP_0 \propto {}_iP_t$
I(W.G.b.) . . .	${}_iP_0 \cdot {}_iQ_0 \propto {}_iP_t \cdot {}_iQ_t$
I(W.G.c.) . . .	${}_iP_{t-q} \cdot {}_iQ_{t-q} \propto {}_iP_t \cdot {}_iQ_t$
I(W.G.a.) . . .	—
I(Laspeyres) . . .	${}_iQ_0 \propto {}_iQ_t$
I(Paasche) . . .	${}_iQ_{t-q} \propto {}_iQ_t$
I(F.)	${}_iQ_0 \propto {}_iQ_{t-q} \propto {}_iQ_t$
I(Agg.a.) . . .	—

using arbitrary year weights, namely, I(W.A.a.), I(W.G.a.) and I(Agg.a.), the assumption so far is that the weighting system is not changed. That is, in spite of the change of the base year from 0 to t, the weights used are still those of the same arbitrary year a. In a later section, there will be a discussion of this particular problem of maintaining continuity when an arbitrary year's weights are changed to those of another arbitrary year, both accompanying and not accompanying a change in base year.

The question naturally arises at this stage: to what extent are any of the conditions likely to be encountered in practice? In planning the compilation of an index number, the problem of continuity, if it is conceded to be an important one, can generally be solved by planning for the switch from the original base to a new base year in such a way that the conditions necessary for the ratio continuity conversion, and relevant to the particular method of construction, are likely to be observed or approximately observed. Thus, if a weighted arithmetic mean is used, having as weights base year values, the switch from the original base year to a new base year can be planned in such a way that the quantities consumed (or produced, or exported, and so on) of the various commodities in the original base year are proportionate to those in the new base year.

In fact, however, many index numbers have had their base year, among other things, changed at regular intervals, and it is important, before using the ratio conversion to obtain a continuous series, to determine the proportionality of the factors relevant to the particular method of construction.

B. *Maintaining Continuity when One or More New Commodities are Substituted for the Equivalent Number of Old Commodities.*

When an index number has been in existence for a period of years it may be found that some old commodities are no longer representative of the group or basket for which the index number was originally compiled, and that some other commodities have come into existence and/or have assumed a greater importance than previously. The index number is accordingly changed to incorporate the new commodities in place of the old ones. The problem is: How to convert the old series before the date of the substitution to the new series with the changed coverage?

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Three points are relevant in discussing this problem, namely:

(*a*) The substitution does not change the general scope and purpose of the index number or there would be no point in obtaining a continuous series.

(*b*) The number of new commodities is the same as that of obsolete commodities.

(*c*) In most cases the new commodities are new only in the sense that they have assumed greater importance; not in the sense of being non-existent before. If, however, the latter is the case, then no data of any kind relevant to the construction of the index number (such as price) would be available before the date of change, and consequently continuity cannot be maintained.*

Substitution is in fact the process of deleting certain commodities and of adding a corresponding number of other (similar) commodities. A change involving substitution is therefore the same as a change involving addition and deletion of commodities simultaneously. As such, the problem is discussed under the next section (Section Two), where the maintenance of continuity is considered in connection with simultaneous changes in several aspects of index number compilation.

c. *Maintaining Continuity when One or More Commodities are Added to the Original Coverage of Commodities*

The index number, I, on year 0 = 100, has been constructed based on a coverage of N commodities up to year T. In year T, n commodities are to be added to this list.† For convenience sake these n commodities will be denoted by symbols $N+1, N+2, \ldots, N+n$. It is required to consider for the various methods of index number construction how to maintain continuity of the series from year 0 on to incorporate the new commodities, and without recalculation.

The discussion will be divided into the following sections, using the notations:

$$i = 1, 2, \ldots, N;\ i' = 1, 2, \ldots, N, N+1, \ldots, N+n;$$

$$j = N+1, N+2, \ldots, N+n;$$

and

$$t = \text{year } 0, 1, 2, \ldots, T-1.$$

(1) The unweighted arithmetic mean method. Here the index on the old coverage is

$$I_{t0} = \frac{100}{N} \sum_i \frac{{}_iP_t}{{}_iP_0},$$

and that on the new coverage is

$$I'_{t0} = \frac{100}{N+n} \sum_{i'} \frac{{}_{i'}P_t}{{}_{i'}P_0}$$

$$= \frac{1}{N+n} \left(N \cdot I_{t0} + 100 \sum_j \frac{{}_jP_t}{{}_jP_0} \right).$$

No recalculation is necessary. The new index number for year t ($t = 0, 1, 2, \ldots, T-1$) can be obtained without difficulty from the old index for the same year by simply applying to the old index the adjustment given in the second term in the above equation, and the data required relate only to the added commodities. If N denotes a large number of commodities, say 100, and if the number n of commodities involved in the change is small, say 5,‡ the saving in labour both in the amassing of past data and in computation is readily appreciated.

* Not directly, that is. To what extent continuity can be achieved by assuming proportionality of the new index number to the old will be discussed in the next article.

† Without, it is to be stressed, altering the original general scope and purpose of the index number. It is also to be stressed that the added commodities here are regarded as new only in the sense that they have assumed sufficient importance to be included in the commodity list for the construction of the index. Their existence in years previous to T is assumed, and therefore data of their prices are available from year 0 onwards. If, however, they are new in that they did not exist before, or that their price data prior to year T are not obtainable the solutions given in this section are obviously not applicable, and reference is made to the next article for a different approach.

‡ This is, in fact, visualized as typical of the situation under discussion. If a large number of commodities are involved in the change, the index number may well assume a new and completely different lease of life, and it may not be worth while to consider the maintenance of continuity.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(2) The unweighted geometric mean index number. A similar relationship can be derived in this case; the new index can be expressed in terms of the old and the added price relatives in the following form:

$$I'_{t0} = \sqrt[(N+n)]{(I_{t0})^N \cdot \prod_j \frac{{}_jP_t}{{}_jP_0}}.$$

It is possible to draw the same conclusions as for the unweighted arithmetic mean index number.

(3) The weighted average of relatives. This sub-section deals with the relatively simple problem of maintaining continuity when commodities are added to the original list of commodities which, because neither the base year 0 nor the arbitrary year a is involved in the change, will continue to be allocated the same value weights (or "basket of goods" weights in the case of ratio-of-aggregatives index numbers) as before the change was introduced.

The problem of maintaining continuity can be dealt with systematically, depending on the weighting.

(i) Weighting by base year weights. Here we deal with the arithmetic mean method first. We have:

$$I_{t0} = \frac{100}{\sum {}_iP_0 \cdot {}_iQ_0} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_0 \cdot {}_iQ_0,$$

and the new index number is:

$$I'_{t0} = \frac{100}{\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0} \sum_{i'} \frac{{}_{i'}P_t}{{}_{i'}P_0} \cdot {}_{i'}P_0 \cdot {}_{i'}Q_0$$

$$= \frac{1}{\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0} \left[I_{t0} \cdot \sum_i {}_iP_0 \cdot {}_iQ_0 + 100 \left(\sum_j \frac{{}_jP_t}{{}_jP_0} \cdot {}_jP_0 \cdot {}_jQ_0 \right) \right].$$

The base year values for all commodities including the added commodities are required and, if available, are constant for every year. For every year t only the price relatives for the added commodities will be needed.

A similar result for the weighted geometric mean method of construction is obtained, as follows:

$$I_{t0} = \sum_{i'} {}_iP_0 \cdot {}_iQ_0 \sqrt{\prod_i \left(\frac{{}_iP_t}{{}_iP_0} \right)^{{}_iP_0 \cdot {}_iQ_0}},$$

$$I'_{t0} = \sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0 \sqrt{(I_{t0})^{\sum_i {}_iP_0 \cdot {}_iQ_0} \cdot \prod_j \left(\frac{{}_jP_t}{{}_jP_0} \right)^{{}_jP_0 \cdot {}_jQ_0}},$$

with again the same implication and conclusions as before.

(ii) Weighting by arbitrary year weights. It can be readily verified that the relationships for both the arithmetic mean and the geometric mean, using arbitrary year value weights, are the same as those using base year value weights, except that where values $P.Q$ are concerned, the suffix throughout will be changed to a denoting arbitrary year instead of 0. Here the data required in addition to the price relatives of the added commodities for every year t are the arbitrary year values for all commodities, and these will be constant.

(iii) Weighting by current year weights. In this case no simplification is possible; the data for every year t in respect of values and of price relatives for the added commodities are necessary. It is clear by now that this is so with all index numbers entailing weighting by current year weights.*

We shall now consider the case of weighting where the weights are given in percentage forms. Some index numbers, although based on base year or arbitrary year value weights, have these weights expressed in percentages. When commodities are added to the original list, these percentages will be altered to make allowance for the relative importance of the new commodities. Since in this section we are not concerned with any change in the weights of the original commodities, if some extra commodities are added with their corresponding weights, the weights of

* See, however, the discussion in the second paper.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the original commodities would be redistributed proportionately. The general solution of the problem may prove to be of interest.

The original commodities i ($i = 1, 2, \ldots, N$) have weights W_i say, and the revised commodities i' ($i' = 1, 2, \ldots, N, N+1, \ldots, N+n$) have weights

$$w_{i'} \quad (\sum_{i'} w_{i'} = \sum_{i} W_i = 100).$$

If we introduce the condition that the weights w_i of the original commodities i are distributed in the same proportion as their former weights W_i, then of course each w_i will be smaller than the corresponding W_i, but proportionately so—i.e.

$$\frac{w_i}{W_i} = a < 1 \text{ for every } i.$$

Under this condition, I'_{t0} can be expressed in terms of I_{t0} and the price relatives of the added commodities just as simply as before. For the weighted arithmetic mean,

$$I'_{t0} = a \cdot I_{t0} + \sum_{j} \frac{{}_jP_t}{{}_jP_0} \cdot w_j,$$

and for the weighted geometric mean,

$$I'_{t0} = (I_{t0})^a \cdot \sqrt[100]{\prod_{j} \left(\frac{{}_jP_t}{{}_jP_0}\right)^{w_j}}$$

where $j = N+1, N+2, \ldots, N+n$ as before, and therefore denotes the added commodities.

(4) Ratio of aggregatives methods of construction. As before, this discussion will be divided into the following parts:

(i) Involving base year quantities, I(Laspeyres). In this case

$$I_{t0} = \frac{\sum_{i} {}_iP_t \cdot {}_iQ_0}{\sum_{i} {}_iP_0 \cdot {}_iQ_0} \cdot 100 = \frac{A}{B} \cdot 100 \text{ say,}$$

$$I'_{t0} = \frac{\sum_{i'} {}_{i'}P_t \cdot {}_{i'}Q_0}{\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0} \cdot 100 = \frac{A+C}{B+D} \cdot 100 \text{ say,}$$

where

$$C = \sum_{j} {}_jP_t \cdot {}_jQ_0 \text{ and } D = \sum {}_jP_0 \cdot {}_jQ_0.$$

Therefore

$$I'_{t0} = \left(\frac{\frac{1}{100} \cdot B \cdot I_{t0} + C}{B+D}\right) \cdot 100.$$

In this equation, C and D involve the added commodities; B involves only the data for the base year 0, and is constant for every year t.

It is easy to verify that the above relationship is identical with that derived for the weighted arithmetic mean using base year value weights.

(ii) Using arbitrary year quantities, I(Agg.a.). Once again the same relationship can be obtained as under (i), the suffix for the Q's being changed to a instead of 0.

(iii) Using current year quantities, I(Paasche). As for the weighted averages using current year value weights, no simple solution is possible.

D. *Maintaining Continuity when One or More Commodities are Deleted from the Original List of Commodities*

The problem here is the converse of that of the previous section, and the solution will be found to be very similar.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We first state the problem symbolically, as follows: The index number I_{t0} has been constructed on a coverage of N commodities ($i = 1, 2, \ldots, N$). Of these, n are now to be deleted, and for convenience sake it will be assumed that the first n of the N commodities are so affected. We shall use the symbol k to denote these ($k = 1, 2, \ldots, n$). A new index number I'_{t0} is required based on the new coverage $i' = n + 1, n + 2, \ldots, N$. The chief consideration is, as before, the conversion from I_{t0} to I'_{t0} without too much of the bother involved in a reassembly of the past data and in recalculation. We shall again follow the lines of the previous sections.

(1) The unweighted arithmetic mean of relatives. We have:

$$I_{t0} = \frac{100}{N} \sum_i \frac{{}_iP_t}{{}_iP_0}$$

$$I'_{t0} = \frac{100}{N-n} \sum_{i'} \frac{{}_{i'}P_t}{{}_{i'}P_0}$$

$$= \frac{1}{N-n}\left[N \cdot I_{t0} - 100 \sum_k \frac{{}_kP_t}{{}_kP_0}\right],$$

the only data required involving the affected commodities k.

(2) The unweighted geometric mean of relatives. The solution is just as easily found, and can readily be verified to be as follows:

$$I'_{t0} = \sqrt[(N-n)]{(I_{t0})^N / \prod_k \frac{{}_kP_t}{{}_kP_0}}.$$

An example illustrates how the unweighted arithmetic mean and the unweighted geometric mean method of construction can easily be made continuous both for addition of commodities and for deletion of commodities. Suppose an old index number, I, has been constructed on a coverage of 100 commodities, and suppose that (a) three commodities, A′, B′ and C′ are to be added to this coverage, and (b) three commodities, A, B and C are to be deleted. The necessary data are given in Table 3:

TABLE 3

Continuity Conversion in the Case of Addition and of Deletion of Commodities. The Data given here Relate only to the Commodities Involved in the Change.

Commodity Deleted	P_0	P_t	$\frac{P_t}{P_0} \cdot 100$	*Commodity Added*	P_0	P_t	$\frac{P_t}{P_0} \cdot 100$
A	1/–	1/6	150	A′	1/3	1/6	120
B	3/–	1/6	50	B′	2/–	2/–	100
C	4/–	5/–	125	C′	7/6	15/–	200

(1) With the unweighted arithmetic mean method of construction, if for the year t the old index number, I_{t0}, had been calculated to be 150·7, then:

(a) For the addition of commodities, the new index is

$$I'_{t0} = \frac{100}{103} \times 150{\cdot}7 + \frac{1}{103}(120 + 100 + 200)$$

$$= \underline{150{\cdot}4.}$$

(*b*) For deletion of commodities, we obtain the new index as

$$I'_{t0} = \frac{100}{97} \times 150{\cdot}7 - \frac{1}{97}(150 + 50 + 125)$$

$$= \underline{152{\cdot}0.}$$

(2) With the unweighted geometric mean method of construction, if, for the same year t say, the old index had been calculated to be 138·6, then

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(*a*) For addition of commodities, the new index is

$$I'_{t0} = \sqrt[103]{138 \cdot 6^{100} \times 120 \times 100 \times 200}$$
$$= \underline{138 \cdot 5}.$$

(*b*) For deletion of commodities, the new index is

$$I'_{t0} = \sqrt[103]{\frac{(138 \cdot 6)^{100}}{150 \times 50 \times 125}}$$
$$= \underline{140 \cdot 2}.$$

(3) Weighted average of relatives. The general results of this and of the following sub-section, dealing with the aggregative methods, can be stated here. The conversion can again be carried out, but slightly more data than for the unweighted average methods will be required for weighting by base year and by arbitrary year values or quantities, while for current year weighting no simple conversion is possible. The additional data required are in fact the weights for all the commodities, but these will again be constant for every year t, ($t = 0, 1, 2, \ldots, T-1$).

(i) Weighting by base year values, for I(W.A.b.) and I(W.G.b.). For the weighted arithmetic mean of relatives, it can be verified that the new index

$$I'_{t0} = \frac{1}{\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0} \left[I_{t0} \cdot \sum_i {}_iP_0 \cdot {}_iQ_0 - 100 \cdot \sum_k \frac{{}_kP_t}{{}_kP_0} \cdot {}_kP_0 \cdot {}_kQ_0 \right].$$

And in the case of the weighted geometric mean of relatives:

$$I'_{t0} = \sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0 \sqrt{(I_{t0})^{\sum_i {}_iP_0 \cdot {}_iQ_0} / \prod_k \left(\frac{{}_kP_t}{{}_kP_0}\right)^{{}_kP_0 \cdot {}_kQ_0}}.$$

(ii) Weighting by arbitrary year values, for I(W.A.a.) and I(W.G.a.). The results for these two index numbers are once again the same as those of the last section with all the value weights altered to arbitrary year instead of base year value weights, i.e., instead of P_0Q_0, we have P_aQ_a throughout.

(iii) Weighting by current year values. No solution of a simple nature is possible in this case, as before.

The solution of the problem in the event of percentage weighting may again prove to be interesting, and is shown as follows. The original commodities i have percentage weights W_i, and the new commodities i' have weights $w_{i'}$. If the weights of the unaffected commodities are distributed proportionately to their weights before the change, then each w_i will be proportionately larger than the corresponding W_i—i.e.,

$$\frac{w_i}{W_i} = b \text{ (say)} > 1 \text{ for every } i.$$

The following relationships can be derived: For the weighted arithmetic mean method of construction,

$$I'_{t0} = b \cdot I_{t0} - \sum_k \frac{{}_kP_t}{{}_kP_0} \cdot w_k$$

and for the weighted geometric mean method of construction,

$$I'_{t0} = (I_{t0})^b \Big/ \sqrt[100]{\prod_k \left(\frac{{}_kP_t}{{}_kP_0}\right)^{w_k}}$$

(4) Ratio of aggregatives methods of construction. The index number using current year quantities again has no simple solution. The index numbers using base year quantities and arbitrary year quantities, I(Laspeyres) and I(Agg.a.) respectively, yield the following relationships:

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(i) Base year quantities—

$$I'_{t0} = \frac{\frac{I_{t0}}{100} \cdot \sum_i {}_iP_0 \cdot {}_iQ_0 - \sum_k {}_kP_t \cdot {}_kQ_0}{\sum_i {}_iP_0 \cdot {}_iQ_0 - \sum_k {}_kP_0 \cdot {}_kQ_0},$$

(ii) Arbitrary year quantities—

$$I'_{t0} = \frac{\frac{I_{t0}}{100} \cdot \sum_i {}_iP_0 \cdot {}_iQ_a - \sum_k {}_kP_t \cdot {}_kQ_a}{\sum_i {}_iP_0 \cdot {}_iQ_a - \sum_k {}_kP_0 \cdot {}_kQ_a}.$$

The following examples illustrate the use of some of the results derived so far with regard to index numbers involving some form of weighting. The data, which are fictitious, are shown in Table 4, and the illustrations will be confined to the results obtained for those index numbers using arbitrary year weighting.

TABLE 4

Fictitious Data to Illustrate the Application of the Results of the Relationships for Maintaining Continuity in the Event of Addition or Deletion of Commodities for Index Numbers Involving Arbitrary Year Value or Quantity Weights. (*See text for the actual working.*)

	1	2	3	4	5	6	7	8	9
Commodity	P_0	P_t	$\frac{P_t}{P_0}.100$	P_a	Q_a	P_aQ_a	$\frac{P_t}{P_0}.P_aQ_a.100$	P_tQ_a	P_0Q_a
A	2/–	2/6	125	2/–	25	50/–	6,250	62·5/–	50/–
B	1/3	2/6	200	1/6	44	66/–	13,200	110/–	55/–
C	2/–	1/6	75	1/6	38	57/–	4,275	57/–	76/–
D	2/–	3/–	150	2/6	60	150/–	22,500	180/–	120/–
E	1/–	2/6	250	1/3	72	90/–	22,500	180/–	72/–
F	3/–	4/6	150	3/6	38	123/–	18,450	171/–	114/–
G	1/–	1/3	125	1/–	32	32/–	4,000	40/–	32/–
H	1/–	2/–	200	2/–	19	38/–	7,600	38/–	19/–
Sum A to H . .	—	—	—	—	—	606/–			
I	2/–	3/6	175	1/6	28	42/–	7,350	98/–	56/–
J	1/–	2/9	275	2/6	62	155/–	42,625	170·5/–	62/–
Sum A to J . .	—	—	—	—	—	803/–	148,750	1,107/–	656/–
Sum I and J . .	—	—	—	—	—	—	49,975	268·5/–	118/–
K	1/–	1/3	125	2/–	15	30/–	3,750	18·75/–	15/–
L	1/6	2/3	150	1/6	46	69/–	10,350	103·5/–	69/–
Sum A to L . .	—	—	—	—	—	902/–			
Sum K and L . .	—	—	—	—	—	—	14,100	122·25/–	84/–

In these examples, t refers to one of the years from 1 to $(T - 1)$. All the old index numbers are assumed to have been compiled on the basis of the ten commodities A to J, and the new index numbers are to be constructed (a) in the case of addition of commodities, with the two commodities K and L added, and (b) in the case of deletion of commodities, with the two commodities I and J deleted.

Although the price data for the unaffected commodities (A to H) are given for year t, they are given purely for the purpose of calculating the I_{t0}. In practice what is visualized is that these data are either unobtainable or difficult to obtain. The value weights P_aQ_a are, however, assumed to be known or available for all commodities, whether affected or otherwise, and the price data are taken to be available only for the affected commodities, I to L. Referring to Table 4, therefore, the following data are to be used for calculating the new index numbers: (i) Columns 1 and 2 (the price data) for commodities I to L only; (ii) Columns 4 and 5, and therefore 6 (the weights) for all commodities A to L; and (iii) Columns 3, 7, 8 and 9 (involving the derived data) for commodities I to L only.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(1) Addition of commodities. The old indices are computed on the ten commodities A to J (i = A, B, . . . , J), and the new indices involve the addition of the two commodities K and L (i' = A, B, . . . , L; j = K, L).

(a) The weighted arithmetic mean method of construction. The old index is

$$I_{t0} = \frac{100}{\sum_i {}_iP_a \cdot {}_iQ_a} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_a \cdot {}_iQ_a$$

$$= \frac{148{,}750}{803} = 185{\cdot}2.$$

The new index, based on the twelve commodities A to L is

$$I'_{t0} = \frac{1}{\sum_{i'} {}_{i'}P_a \cdot {}_{i'}Q_a} \left[I_{t0} \cdot \sum_i {}_iP_a \cdot {}_iQ_a + \left(\sum_j \frac{{}_jP_t}{{}_jP_0} \cdot {}_jP_a \cdot {}_jQ_a \right) 100 \right]$$

$$= \frac{1}{902} [(185{\cdot}2 \times 803) + (14{,}100)]$$

$$= \underline{180{\cdot}5.}$$

(b) The weighted geometric mean method. The old index is

$$I_{t0} = \sum_i {}_iP_a \cdot {}_iQ_a \sqrt{\prod_i \left(\frac{{}_iP_t}{{}_iP_0} \right)^{{}_iP_a \cdot {}_iQ_a}}$$

$$= 174{\cdot}6.$$

The new index is

$$I'_{t0} = \sqrt[902]{(174{\cdot}6)^{803} \cdot (125)^{30} \cdot (150)^{69}}$$

$$= \underline{170{\cdot}6.}$$

(c) The ratio of aggregatives of form I(Agg.a). The old index is

$$I_{t0} = \frac{\sum_i {}_iP_t \cdot {}_iQ_a}{\sum_i {}_iP_0 \cdot {}_iQ_a} \cdot 100$$

$$= 168{\cdot}75.$$

The new index is

$$I'_{t0} = \frac{\frac{168{\cdot}75}{100} \cdot 656 + 122{\cdot}25}{656 + 84}$$

$$= \underline{166{\cdot}1.}$$

(2) Deletion of commodities. The original index number is again based on the ten commodities A to J, and the new index number involves the deletion of the two commodities I and J. With reference to the symbols used above, i = A, B, . . . , J; i' = A, B, . . . , H; and k = I, J.

(a) The weighted arithmetic mean method of construction:

$$I_{t0} = 185{\cdot}2.$$

$$I'_{t0} = \frac{1}{606} [(185{\cdot}2)(803) - (7{,}350 + 42{,}625)]$$

$$= \underline{162{\cdot}9.}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(b) The weighted geometric mean method of construction:

$$I_{t0} = 174 \cdot 6.$$

$$I'_{t0} = \sqrt[606]{(174 \cdot 6)^{803}/(175)^{42} \cdot (275)^{155}}$$

$$= \underline{155 \cdot 4}.$$

(c) The ratio of aggregatives method of construction:

$$I_{t0} = 168 \cdot 75.$$

$$I'_{t0} = \frac{\frac{168 \cdot 75}{100} \times 656 - 268 \cdot 5}{656 - 118}$$

$$= \underline{155.9}.$$

E. *Maintaining Continuity when the Weighting System is Altered*

This problem is also commonly encountered. After an index number has been in existence for a period of time, the original weighting system may no longer represent the relative importance of the commodities in respect of the purpose of the index. A new weighting system is therefore set up, and it is our concern to study how to maintain the continuity of the series before and after the change. It goes without saying that all the unweighted average methods of construction would be excluded from this discussion.

The index numbers involving base year weights are the easiest to deal with in this connection, for with all such index numbers a change of the base year automatically means a change of the weighting system from that of the original base year to that of the new base year. The maintenance of continuity with a change in the base year has been discussed in the first section of this thesis, where the conditions for the application of the simple ratio continuity conversion are presented in Table 2.

Unfortunately, this apparently simple solution of the problem is unsatisfactory, for in the section dealing with changes in the base year the main emphasis is on the base year, the weighting system being subsidiary. The view taken there was something like this: to maintain the continuity of the series in the event of changes in the base year, the ratio continuity conversion can be used, but only under certain conditions in the case of most of the index numbers. The snag is that these conditions are in some way or other connected with the weighting system, which is the main consideration in this present section. Generally the weighting system is changed because the proportionate importance of each commoditity represented by its weight is believed to have changed radically. We cannot therefore assume proportionality in the weighting system, which is in effect what Table 2 requires in the majority of cases.

The solution of the problem will be postponed to my second paper, where the maintenance of continuity will be studied from a different approach from the direct one adopted here.

The arbitrary year weighted index numbers are in effect subject to the same problems and the same difficulties, and will also be dealt with later.

Finally, with regard to current-year weighted index numbers, by the very definition of such index numbers, the weights are required to be changed every year. Hence, within the scope of their interpretations, other things being equal, these index numbers maintain their continuity automatically as far as weight changes are concerned.

SECTION TWO

Maintaining Continuity where there are Two or More Simultaneous Changes in Index Number Compilation

Introduction

So far, the discussion has been concerned with cases where one alteration in the compilation of the index number takes place. The more usual practice is to change more than one aspect of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the compilation, and it can be readily appreciated that the problem of maintaining continuity becomes more and more complicated the more changes are introduced. In this section we shall attempt to consider such more general cases, but before doing so we must systematize the schemes and the notations so that when any situation becomes complicated, we shall at any rate not have to be further confused by non-uniform notations and symbols.

In this section, therefore, we shall assume that the original list of commodities consists of N items, $i = 1, 2, \ldots, N$. Addition will involve n commodities, and it will be assumed that these n commodities will be added at the end of the original list. Using k to denote the added commodities, $k = N + 1, N + 2, \ldots, N + n$. The deleted commodities will be assumed to affect the last p of the original N commodities, and will be denoted by h, $h = N - p + 1, N - p + 2, \ldots, N$. After addition and deletion have taken place, the new list of commodities will consist of $(N + n - p)$ items, and will be denoted by i', where

$$i' = 1, 2, \ldots, N - p, N + 1, \ldots, N + n.$$

The substitution of commodities, as mentioned in the last section, can be regarded as a simultaneous change where n of the original commodities, say, are replaced by an equivalent number of similar commodities, so that, in the above system of notations, $p = n$. The new list of commodities will in this case be N; h will range from $N - n + 1$ to N, and k from $N + 1$ to $N + n$. Although this is not important, it may clarify the matter to regard the commodity $N - n + 1$ as being replaced by commodity $N + 1$, $N - n + 2$ by $N + 2$, etc., and N by $N + n$.

The important proviso in any of the changes introduced is that such changes do not alter the original general purpose or general scope of the index numbers. If they do, no reason can arise for wishing to maintain the continuity of two dissimilar index numbers.

A. *Simultaneous Changes Involving Addition and Deletion of Commodities; the Problem of Substitution*

The original index number I_{t0} (year 0 = 100) was based on N commodities, and the new index number I'_{t0} (Year 0 = 100) is now based on $N + n - p$ commodities. The discussion will again be subdivided into the following headings:

(1) The unweighted arithmetic mean of relatives. The solution in this case is an extension of the results obtained in the last chapter. The old index is—

$$I_{t0} = \frac{100}{N} \sum_i \frac{{}_iP_t}{{}_iP_0},$$

and the new index is

$$I'_{t0} = \frac{100}{N + n - p} \sum_{i'} \frac{{}_{i'}P_t}{{}_{i'}P_0}$$

$$= \frac{1}{N + n - p} \left\{ N \cdot I_{t0} + 100 \left[\sum_k \frac{{}_kP_t}{{}_kP_0} - \sum_h \frac{{}_hP_t}{{}_hP_0} \right] \right\}.$$

(2) The unweighted geometric mean of relatives. Similarly,

$$I'_{t0} = \sqrt[(N+n-p)]{(I_{t0})^N \cdot \frac{\prod_k ({}_kP_t/{}_kP_0)}{\prod_h ({}_hP_t/{}_hP_0)}}.$$

In both the above cases, only the relatives of the affected commodities will be required in the calculation of the new index number, and no wholesale recalculation is necessary.

To illustrate the method, we shall use an example similar to the one used in the relevant sections in the last chapter. Suppose an index number computed originally on a coverage of 100 commodities; suppose further that two commodities, D and E, are to be deleted, and that three commodities, X, Y and Z are to be added, so that the new coverage will consist of

$$(100 + 3 - 2) = 101 \text{ commodities.}$$

The relevant data for the affected commodities in year t and year 0 are given in Table 5.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(1) With the unweighted arithmetic mean method of construction,

$$I_{t0} = 150{\cdot}7 \text{ (as before)},$$

$$I'_{t0} = \frac{1}{101}\,[(150{\cdot}7 \times 100) + (125 + 120 + 200) - (100 + 75)]$$

$$= \underline{151{\cdot}9}.$$

TABLE 5

Data to Illustrate the Relationships between the New Index Number and the Old Index Number in the Event of Changes due to Addition and Deletion of Commodities Happening Simultaneously. (See text for explanaton.)

Commodity	P_t	P_0	$\frac{P_t}{P_0} \cdot 100$
D	2/–	2/–	100
E	4/–	3/–	75
X	4/–	5/–	125
Y	1/3	1/6	120
Z	2/–	4/–	200

(2) With the unweighted geometric mean method of construction,

$$I_{t0} = 138{\cdot}6 \text{ (as before)},$$

$$I'_{t0} = \sqrt[101]{(138{\cdot}6)^{100} \cdot \frac{125 \times 120 \times 200}{100 \times 75}}$$

$$= \underline{140{\cdot}1}.$$

(3) Weighted average of relatives. Current weighted index numbers are, as before, not simple to convert for continuity. Base-weighted and arbitrary-year-weighted index numbers can, however, be so converted, the additional data required, other than those needed in the case of the unweighted index numbers, being the base-year or arbitrary-year value weights of the affected commodities.

For the weighted arithmetic mean index numbers using base-year and arbitrary-year value weights respectively, i.e., for I(W.A.b.) and I(W.A.a.) respectively, very similar relationships apply. Thus, for I(W.A.b.),

$$I'_{t0} = \frac{1}{\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0}\left[I_{t0} \cdot \sum_i {}_iP_0 \cdot {}_iQ_0 + 100\left(\sum_k \frac{{}_kP_t}{{}_kP_0} \cdot {}_kP_0 \cdot {}_kQ_0 - \sum_h \frac{{}_hP_t}{{}_hP_0} \cdot {}_hP_0 \cdot {}_hQ_0\right)\right]$$

and exactly the same relationship for I(W.A.a.), except that all through the values should refer to arbitrary year values instead of base year values, and should read $P_a \cdot Q_a$ instead of $P_0 \cdot Q_0$ with the prefixes unchanged.

Similarly the relationships for the weighted geometric means of relatives can be derived. The relationship for the index number weighted by base-year value weights is shown here, i.e., for I(W.G.b.); that for the index number weighted by arbitrary year value weights, I(W.G.a.) is again the same except for the values which should refer to the arbitrary year a instead of the base year 0. The relationship for the base-weighted index number is therefore as follows:

$$I'_{t0} = \sqrt[\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0]{(I_{t0})^{\sum_i {}_iP_0 \cdot {}_iQ_0} \cdot \frac{\prod_k \left({}_kP_t / {}_kP_0\right)^{{}_kP_0 \cdot {}_kQ_0}}{\prod_h \left({}_hP_t / {}_hP_0\right)^{{}_hP_0 \cdot {}_hQ_0}}}.$$

(4) Ratio of aggregatives index numbers. Again no conversion is possible for the index number based on the current "basket of goods". In respect of index numbers based on the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

base-year "basket of goods" and on the arbitrary year "basket of goods", the relationships of the revised and the original indices are: first, for the base year "basket of goods",

$$I'_{t0} = \frac{100\left[\frac{I_{t0}}{100} \cdot \sum_i {}_iP_0 \cdot {}_iQ_0 + \sum_k {}_kP_t \cdot {}_kQ_0 - \sum_h {}_hP_t \cdot {}_hQ_0\right]}{\sum_i {}_iP_0 \cdot {}_iQ_0 + \sum_k {}_kP_0 \cdot {}_kQ_0 - \sum_h {}_hP_0 \cdot {}_hQ_0}$$

and for the arbitrary year "basket of goods", the same relationship holds as above, except that all the Q_0's are changed to Q_a for the arbitrary year.

In all the results derived in this section, the relationships are not so complicated as they appear, if it is remembered that: (i) the number of commodities which are affected and for which data are required for every year t is usually very small compared with the whole list of commodities, and (ii) the data for the original complete list of commodities are required only for year 0 or year a and are constant for all years t.

The data of Table 4 of the previous section will now be used to illustrate the application of the relationships derived with regard to arbitrary year weighted index numbers. It will be assumed that the original commodities consisted of 10 items A to J. The changes involved in the revision of this original list of 10 commodities are: (i) Commodities K and L are to be added; and (ii) Commodity J is to be deleted. i' will here denote the commodities A to I plus K and L, so that $\sum_{i'} {}_{i'}P_0 \cdot {}_{i'}Q_0 = 747$.

First of all, with the weighted arithmetic mean method of construction,

$$I_{t0} = 185{\cdot}2 \text{ as derived before,}$$

$$I'_{t0} = \frac{1}{747}\left[185{\cdot}2\,(803) + (3750 + 10350 - 42625)\right]$$

$$= \underline{160{\cdot}9.}$$

With the weighted geometric mean method of construction,

$$I_{t0} = 174{\cdot}6,$$

$$I'_{t0} = \sqrt[747]{(174{\cdot}6)^{803} \cdot \frac{125^{30} \cdot 150^{69}}{275^{155}}}$$

$$= \underline{154{\cdot}5.}$$

Finally, with the aggregative-ratio index number,

$$I_{t0} = 168{\cdot}75,$$

$$I'_{t0} = \frac{100\left(\frac{168{\cdot}75}{100} \cdot 656 + 18{\cdot}75 + 103{\cdot}5 - 170{\cdot}5\right.}{656 + 15 + 69 - 62}$$

$$= \underline{156{\cdot}2.}$$

Substitution is to be regarded as a special case of this problem of simultaneous changes involving addition and deletion of commodities. The number of commodities added (n) is the same as the number of commodities deleted, and in the identities developed in this sub-section, whenever substitution alone is considered, the new list of commodities will consist of the same number (N) as the old list; k will denote the commodities $N+1, N+2, \ldots, N+n$, which are substituted for the commodities denoted by h, $h = N-n+1, N-n+2, \ldots, N$.

If the new index number involves simultaneous changes of substitution, addition and deletion of commodities, an obvious extension of the solutions considered in this sub-section would lead us to include the substituting commodities with the added commodities and the substituted commodities with the deleted commodities, and to regard the problem as one involving addition and deletion simultaneously without further complication.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B. *Changes Due to Substitution, Addition and Deletion of Commodities Simultaneously with a Shift in the Base Year*

The index number, I, has been constructed on: (i) base year $0 = 100$; and (ii) a coverage of N commodities ($i = 1, 2, \ldots, N$). In year t both these aspects of construction are to be changed as follows: (i) The base year is to be shifted to $t = 100$; and (ii) The first m of the N commodities to be substituted by the same number of similar commodities, the last p of the N commodities to be deleted, and n other commodities to be added, so that the new list of commodities will, as in the last section, consist of $(N + n - p)$ commodities, ($i' = 1', 2', \ldots, m'$, $m + 1, \ldots N - p, N + 1, \ldots, N + n$ where $1', 2', \ldots, m'$ are substituted for $1, 2, \ldots, m$ respectively). The problem: To consider how to maintain the continuity of the series of index numbers from year 0 beyond year t on the new base year $t = 100$ and incorporating the new coverage of commodities.

Following the line of the first section of the last chapter dealing with a simple shift of the base year, we shall consider the index number for year $t - q$ say, and study the problem of converting the index for this year on base 0 to that of the same year on base t, only here there is the additional matter of commodity coverage. In other words, we shall study the conversion of the index number $I'_{(t-q)0}$ to the index number $I'_{(t-q)t}$, $q = 1, 2, \ldots, t$.

In the present discussion, it will be convenient to separate the situation into two general kinds: first, consisting of those index numbers for which a continuous series can be obtained by the use of the simple ratio conversion procedure when only the base year is changed; second, consisting of those index numbers for which this procedure can be used only under certain conditions.

(a) Index numbers which, for a change in the base year, maintain continuity by simple ratio conversion. A reference to Table 2 shows that three index numbers come under this heading, namely, the simple geometric mean of relatives, the geometric mean method using arbitrary year values as weights, and the aggregative ratio index number based on an arbitrary year "basket of goods".

In the case of these index numbers, if for the commodities i,

$$\frac{I_{(t-q)0}}{I_{t0}} = I_{(t-q)t},$$

then it follows that for commodities i',

$$\frac{I'_{(t-q)0}}{I'_{t0}} = I'_{(t-q)t}.$$

It can then be verified that, using the same notations as in the last section, for the simple unweighted geometric mean method of construction,

$$I'_{(t-q)t} = \sqrt[(N+n-p)]{\left(\frac{I_{(t-q)0}}{I_{t0}}\right)^N \cdot \frac{\prod_{j'} ({}_{j'}P_{t-q}/{}_{j'}P_t)}{\prod ({}_{j}P_{t-q}/{}_{j}P_t)} \cdot \frac{\prod_{k} ({}_{k}P_{t-q}/{}_{k}P_t)}{\prod_{h} ({}_{h}P_{t-q}/{}_{h}P_t)}}, \quad \begin{matrix} j' = 1', 2', \ldots, m', \\ j = 1, 2, \ldots, m. \end{matrix}$$

Similarly, for the weighted geometric mean index number weighted by arbitrary year values,

$$I'_{(t-q)t} = \sqrt[\sum_{i'} {}_{i'}P_a \cdot {}_{i'}Q_a]{\left(\frac{I_{(t-q)0}}{I_{t0}}\right)^{\sum_{i} {}_{i}P_a \cdot {}_{i}Q_a} \cdot \frac{\prod_{j'} \left(\frac{{}_{j'}P_{t-q}}{{}_{j'}P_t}\right)^{{}_{j'}P_a \cdot {}_{j'}Q_a}}{\prod \left(\frac{{}_{j}P_{t-q}}{{}_{j}P_t}\right)^{{}_{j}P_a \cdot {}_{j}Q_a}} \cdot \frac{\prod_{k} \left(\frac{{}_{k}P_{t-q}}{{}_{k}P_t}\right)^{{}_{k}P_a \cdot {}_{k}Q_a}}{\prod_{h} \left(\frac{{}_{h}P_{t-q}}{{}_{h}P_t}\right)^{{}_{h}P_a \cdot {}_{h}Q_a}}}.$$

Finally, the aggregative ratio index number based on an arbitrary year "basket of goods" yields the following relationship:

$$I'_{(t-q)t} = \frac{\frac{I_{(t-q)0}}{100} \cdot \sum_{i} {}_{i}P_a \cdot {}_{i}Q_a + (\sum_{j'} {}_{j'}P_{t-q} \cdot {}_{j'}Q_a - \sum_{j} {}_{j}P_{t-q} \cdot {}_{j}Q_a) + \sum_{k} {}_{k}P_{t-q} \cdot {}_{k}Q_a - \sum_{h} {}_{h}P_{t-q} \cdot {}_{h}Q_a}{\frac{I_{t0}}{100} \cdot \sum_{i} {}_{i}P_a \cdot {}_{i}Q_a + (\sum_{j'} {}_{j'}P_t \cdot {}_{j'}Q_a - \sum_{j} {}_{j}P_t \cdot {}_{j}Q_a) + \sum_{k} {}_{k}P_t \cdot {}_{k}Q_a - \sum_{h} {}_{h}P_t \cdot {}_{h}Q_a}.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To illustrate the above relationships, we shall confine ourselves to the relatively simple case of the unweighted geometric mean method of construction. An index number, using the unweighted geometric mean method of construction, with an original list of 100 commodities, and based on year $0 = 100$, is found to have the values for year $t - q$ and year t as follows: $I_{(t-q)0} = 125{\cdot}6$; $I_{t0} = 138{\cdot}6$. At or subsequent to year t, it is decided to change the base year from 0 to t, and also to substitute three of the original commodities, A, B and C by A′, B′ and C′, delete two of the original commodities, D and E, and add one commodity, The prices in the years $(t - q)$ and t for these affected commodities are given in the following table, where the relatives (P_{t-q}/P_t) are also derived:

Commodity	P_{t-q}	P_t	$\frac{P_{t-q}}{P_t} \cdot 100$
A	1/–	1/–	100
B	2/3	3/–	75
C	3/6	4/–	87·5
A′	1/–	1/3	80
B′	1/–	2/–	50
C′	6/–	7/6	80
D	2/–	2/–	100
E	3/–	4/–	75
Z	2/3	2/6	90

These data are sufficient for the conversion of the index number $I_{(t-q)0}$ to $I'_{(t-q)t}$ on base t and incorporating the several changes due to substitution, deletion and addition of commodities. From the relationship derived above for the unweighted geometric mean method of construction, we obtain:

$$I'_{(t-q)t} = \sqrt[99]{\left(\frac{125{\cdot}6}{138{\cdot}6}\right)^{100} \cdot \left(\frac{80 \times 50 \times 80}{100 \times 75 \times 87{\cdot}5}\right) \cdot \frac{90}{100 \times 75}}$$

$$= \underline{85{\cdot}9}.$$

(b) Index numbers which, when the base year is shifted, maintain continuity by simple ratio conversion but under certain conditions. For all these index numbers, the conditions are similar to those required where only the base year is shifted and no other alterations are introduced, except that instead of commodities i, we now have commodities i'. For example, for the unweighted arithmetic mean method of construction, the condition required for ratio conversion for continuity is now:

$$_{j'}P_0 \propto {}_{i'}P_t \text{ instead of } {}_iP_0 \propto {}_iP_t.$$

If this condition is satisfied, then we obtain

$$I'_{(t-q)t} = \frac{I'_{(t-q)0}}{I'_{t0}}.$$

Two procedures are therefore needed here: (i) The determination of whether or not the prices of the new list of commodities in the original base year are proportionate or approximately proportionate to those in the new base year (or other conditions corresponding to the relevant index numbers). If so, then (ii) the calculation of the index numbers $I'_{(t-q)0}$ and I'_{t0} incorporating the changes due to substitution, addition and deletion of commodities, thereafter applying the ratio conversion to obtain $I'_{(t-q)t}$.

C. *Change in the Weighting System Simultaneously with Changes Due to a Shift in the Base Year and to Substitution, Addition and Deletion of Commodities*

By their very nature the unweighted average methods of construction and the methods involving weighting by current year values or weights will be excluded from the discussion of this section.

In sub-section B, under (a), the two index numbers I(W.G.a.) and I(Agg.a.) were discussed in relation to a shift in the base year simultaneously with changes due to substitution, addition and deletion of commodities. It was found there that there exist relationships for the maintenance

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of continuity of the index numbers. In the derivation of these relationships, the assumption was implicit that the arbitrary year weighting system did not change.

In the same section under (b), the several index numbers involving base year weighting were concerned, among others, in respect of the same changes. The conclusion was reached there that, provided the relevant conditions are satisfied, the ratio conversion procedure can be applied to deal with the shift in the base year after adjustments have been carried out to cover the commodity changes. However, these conditions imply in some way or other proportionality of the weighting system of the original base year to that of the new base year, whereas in fact a weighting system requires to be changed only because this very proportionality is no longer applicable.

The problem is similar to that encountered where only the change in the weighting system is involved, but here it is further complicated by the other changes, namely those concerned with commodity coverage. And as in that case, the solution of the problem cannot be found by the direct approach so far adopted. It will be discussed under a different approach, which will form the subject matter of the second paper.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A Geometrical Derivation of the Analyses of Covariance and Variance

By C. P. Cox

National Institute for Research in Dairying, Shinfield, near Reading

1. Introduction

The generalization of results in three dimensions to illustrate elementary statistical principles is a well-established technique—for example, Champernowne (1946)—and the following note on a vectorial demonstration of the analyses of covariance and variance was prepared some time ago. Its submission is inspired by the President's inaugural address, Pearson (1956).

2. Exposition

With sufficient generality consider a randomized block experiment of n blocks, each containing two treatments; let y_{ij} represent the experimental observation for the ith treatment in the jth block and let x_{ij} represent the corresponding value of the covariate.

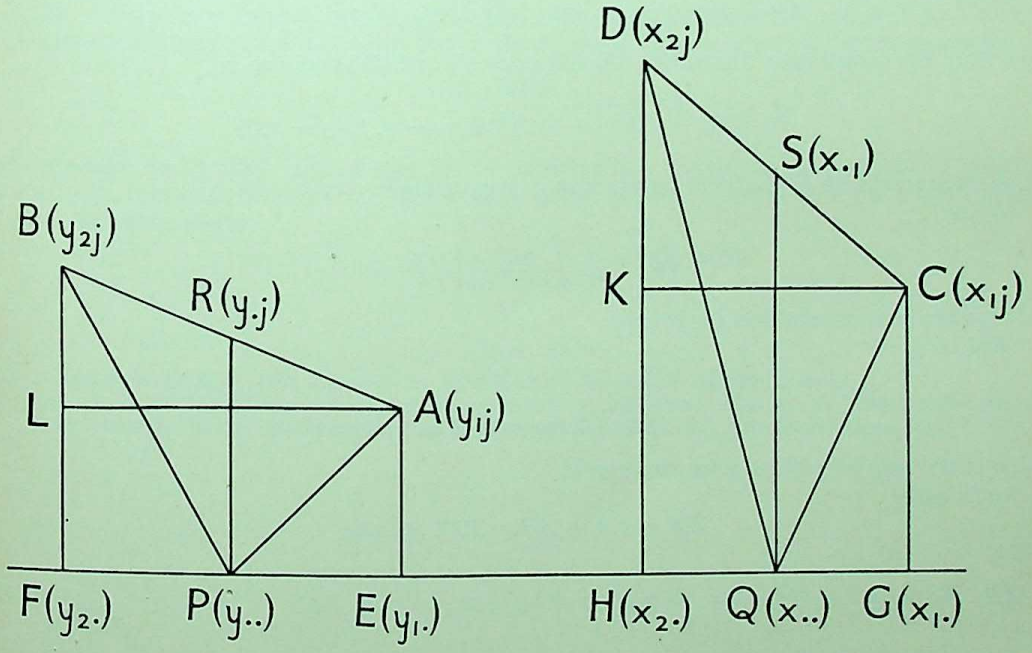


Fig. 1.

In the space of n-dimensions and with reference to Fig. 1, let A be the point whose coordinates y_{1j}, $j = 1, \ldots, n$, are determined by the observations for the first treatment and let B, y_{2j}, $j = 1, \ldots, n$, be similarly determined by the second treatment. The coordinates of R, the midpoint of AB, will then be $(y_{1j} + y_{2j})/2$, $j = 1, \ldots, n$, that is, R is the point $(y_{\cdot 1}, y_{\cdot 2}, \ldots, y_{\cdot n})$ determined by the block means. Let C, D and S be the corresponding points for values of the covariate x, it being observed that AB and CD will not generally lie in the same plane.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Let P and Q be the points $(y_{..}, \ldots, y_{..})$ and $(x_{..}, \ldots, x_{..})$ on the line for which all coordinates are equal and let E, F, G and H be the feet of the perpendiculars to this line from the points A, B, C and D respectively.

If E is the point $(\lambda, \ldots, \lambda)$ we have, for perpendicularity,

$$\sum_{j=1}^{n} (y_{1j} - \lambda)(\lambda - y_{..}) = 0,$$

from which it easily follows that $\lambda = y_{1.}$ and the coordinates of the points F, G and H, as shown, are similarly obtained.

Lastly, let L and K be the feet of the perpendiculars from A to BF and from C to DH.

We then have by vector algebra

$$\begin{aligned} PA \,.\, QC + PB \,.\, QD &= (PR + RA) \,.\, (QS + SC) + (PR + RB) \,.\, (QS + SD) \\ &= 2PR \,.\, QS + RA \,.\, SC + RB \,.\, SD \\ &= 2PR \,.\, QS + \tfrac{1}{2}AB \,.\, CD. \end{aligned}$$

Further,

$$\begin{aligned} AB \,.\, CD &= (AL + LB) \,.\, (CK + KD) \\ &= AL \,.\, CK + LB \,.\, KD, \qquad (1) \end{aligned}$$

since the scalar product of two perpendicular vectors is zero.

Hence

$$PA \,.\, QC + PB \,.\, QD = 2PR \,.\, QS + \tfrac{1}{2}AL \,.\, CK + \tfrac{1}{2}LB \,.\, KD \qquad (2)$$

Now from the coordinates themselves, the left hand side of (2) is seen to be

$$\sum_{j=1}^{n} \{(y_{1j} - y_{..})(x_{1j} - x_{..}) + (y_{2j} - y_{..})(x_{2j} - x_{..})\}$$

which is the total sum of products in the analysis of covariance.

Again,

$$2PR \,.\, QS = 2 \sum_{j=1}^{n} (y_{.j} - y_{..})(x_{.j} - x_{..})$$

which is the sum of products for blocks.

Next,

$$\begin{aligned} \tfrac{1}{2}AL \,.\, CK &= \tfrac{1}{2}EF \,.\, GH \\ &= \tfrac{1}{2}n(y_{2.} - y_{1.})(x_{2.} - x_{1.}) \end{aligned}$$

which is the sum of products for treatments.

And finally

$$LB = LA + AB = 2(PE + AK),$$

so that

$$\tfrac{1}{2}LB \,.\, KD = \tfrac{1}{2} \sum_{j=1}^{n} 4(y_{1.} - y_{..} + y_{.j} - y_{1j})(x_{1.} - x_{..} + x_{.j} - x_{1j}) \qquad (3)$$

$$= \tfrac{1}{2} \sum_{j=1}^{n} 4(y_{..} - y_{2.} + y_{2j} - y_{.j})(x_{..} - x_{2.} + x_{2j} - x_{.j}) \qquad (4)$$

$$\left.\begin{aligned} &= \sum_{j=1}^{n} (y_{1.} - y_{..} + y_{.j} - y_{1j})(x_{1.} - x_{..} + x_{.j} - x_{1j}) \\ &+ \sum_{j=1}^{n} (y_{..} - y_{2.} + y_{2j} - y_{.j})(x_{..} - x_{2.} + x_{2j} - x_{.j}) \end{aligned}\right\} \qquad (5)$$

which is the usual expression for the interaction sum of products. For the case of the two-treatment experiment equations (3) and (4) give explicit expressions for the interaction in terms

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of values for either treatment alone. Alternative forms for the interaction term in this case are also easily obtained. First, since $LB = LA + AB$,

$$\tfrac{1}{2}LB \,.\, KD = \tfrac{1}{2}\sum_{j=1}^{n}(y_{1.} - y_{2.} + y_{2j} - y_{1j})(x_{1.} - x_{2.} + x_{2j} - x_{1j}).$$

And second, since from (1)

$$\tfrac{1}{2}LB \,.\, KD = \tfrac{1}{2}AB \,.\, CD - \text{sum of products for treatments},$$

interaction $S.S. = \tfrac{1}{2}\sum_{j=1}^{n}(y_{2j} - y_{1j})(x_{2j} - x_{1j})$ − sum of products for treatments.

For the analysis of variance it is easily seen that the individual items in the sum of squares partitioning are obtained by superposing the two parts of Fig. 1, i.e. by putting $x_{ij} = y_{ij}$ for all i and j.

The partitioning in cases of more than two treatments can be made in an exactly similar way although the explicit expression for the interaction term will be that corresponding to the form in equation (5).

Interrelations between separate items in the sums of squares can be examined from the figure, it being of particular interest to notice how the interaction term varies with the direction of the line AB. When AB is nearly parallel to EF, that is when the within-block treatment differences $(y_{2j} - y_{1j})$ are nearly equal, the interaction term will be small, and conversely.

It can also be seen in Fig. 1 that the regression of y on x calculated from the interaction term is $\tfrac{1}{2}LB \,.\, KD/\tfrac{1}{2}KD^2 = LB \cos\theta_1/KD$ where θ_1 is the angle between LB and KD, and that the correlation coefficient is $\cos\theta_1$. Also the residual sum of squares allowing for the regression is

$$\tfrac{1}{2}LB^2 - \tfrac{1}{2}\frac{(LB \,.\, KD)^2}{KD^2} = \tfrac{1}{2}LB^2 \sin^2\theta_1,$$

and similarly the residual sum of squares for treatments + interaction is $\tfrac{1}{2}AB^2 \sin^2\theta_2$ where θ_2 is the angle between AB and CD. Thus, since $LB = AB \sin\psi$ where ψ is the angle LAB, the F ratio for treatments is

$$(n-2)\left(\frac{\sin^2\theta_2}{\sin^2\psi\sin^2\theta_1} - 1\right).$$

References

CHAMPERNOWNE, D. G. (1946), *Lectures on Mathematical Statistics.* [Unpublished.]

PEARSON, E. S. (1956), "Some aspects of the geometry of statistics. The use of visual presentation in understanding the theory and application of mathematical statistics", *J.R. Statist. Soc.*, A, **119**, 125–146.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Note on the History of Sampling Methods in Russia

By S. S. Žarković

Various articles have been published on the history of the theory and application of sampling methods. They include: "A review of recent statistical developments in sampling surveys", by F. Yates, published in this *Journal*, Vol. CIX, 1946; "History of the uses of modern sampling procedures", by F. F. Stephen, published in the *Journal of the American Statistical Association*, Vol. 43, 1948; and the "Historical survey of the development of sampling theories and practice", by You Poh Seng, published in this *Journal*, Vol. CXIV, 1951.

Some of these articles have considerable importance because of their systematic approach to the subject. None, however, contains any reference to Russian contributions to the development of sampling methods during the First World War and in the period immediately following it, when Lenin was still alive. I should like to call attention to this fact because, at least as far as I know, it may be said that during those years it was precisely in Russia that the most intensive work in the field of sampling was being carried on.

In this connection the mathematical traditions of Russia should not be overlooked. At the time in question Russian statisticians were aware of the results of Tchebyshev's work. Markov, whose interest in statistical problems is well known, was still living. A. A. Tchuprow was also working in Russia during the First World War, and his contributions to statistical theory are known throughout the world. In addition, a number of highly competent statisticians were active, among them A. Kaufmann, V. Romanovsky, E. Slutsky, B. Iastremsky, Boiarsky, Tchetwerikov, Obukhov, etc. Many of them were young and subsequent events prevented them from becoming known outside Russian borders.

To understand Russian statistics at that time one must remember that the general trend in statistics during the first years after the revolution differed markedly from that apparent later. Judging from publications, it appears that in the earlier period theoretical work was completely free. The problems which Russian post-revolution statistical theory dealt with were those common to many countries at that time and were thus being widely studied. The methods used for solving practical problems were likewise those common to the current statistical methodology.

Another feature of Russian statistics in those years was the interest theoreticians took in practical questions and the effect it had on their intensive application of theory to the solution of practical problems. These Russian statisticians watched the development of statistical theory all over the world, they published translations of the most important foreign contributions and they reviewed for the benefit of their readers all important results, whatever country provided them. This keen activity supplied the base from which they sought solutions to their own practical problems. I would go so far as to say that, in consequence, Russian statistics in the early thirties were on a par with the best in other countries.

A very detailed picture of all aspects of this statistical activity may be obtained by reading *Vestnik Statistiki*, the organ of the Central Statistical Administration, which started publication in 1919. One section of the *Vestnik* was devoted to theory, another to reviews of statistical literature and accounts of the work on which different statistical agencies were engaged. There were also notes on statistically important events and facts, etc. The main source of the information given in this article is, in fact, *Vestnik*.

A study of that publication and of other available sources reveals that sampling methods were in practical use during the First World War and that their application was extensive thereafter. In his book *The Application of Mathematical Statistics to Experimental Work*, Moscow, 1946, V. Romanovsky states that A. A. Tchuprow had designed some sampling surveys during the First World War. In *Vestnik*, Vol. III, 1919, S. S. Kon wrote "On the question of the application of sampling methods in the processing of the agricultural census". In Vol. IV, 1919, there is an article by N. Tchetwerikov, director of the Methodological Division of the Central Statistical Administration, "On sampling methods". In this interesting article, important from the historical point of view, Tchetwerikov says that A. A. Tchuprow read his report "On sampling methods" to the meeting of the Russian research workers in natural sciences and medicine, held in 1900. According to Tchetwerikov, in this report only probability samples are dealt with and the basic theory of surveys is developed.

As far as Tchetwerikov himself is concerned the theory of sampling represents in his opinion an objective procedure based on random selection of sampling units with the known probabilities of selection of each particular unit. The basic feature of sample surveys is not the size of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sample, says Tchetwerikov, as was believed at that time by some Russian statisticians, but the way in which the sampling units are drawn into the sample. The aim of this article is to present the logical bases of sampling methods; with regard to formulae, Tchetwerikov says they can be found in Tchuprow's works. In his attempts to make clear the meaning of sampling methods, Tchetwerikov speaks on simple random samples selected with or without replacement of units, on stratified samples, on cluster sampling, and on systematic selection of sampling units. In his presentation of sampling methods he mentions some problems important even to-day. Thus he says that the precision of estimates does not depend only upon the size of the sample but also on how the units are selected, on the magnitude of the variation of the phenomenon under study and the shape of the distribution. In connection with the latter he calls attention to the need for the new theoretical developments in order to show how the precision of estimates is influenced by distributions deviating from the normal type. This is the field in which new results are expected even to-day. In addition, the possibility of using supplementary information in sample surveys is described, and so the article covers the basic chapters of the modern theory of sampling.

In Vol. VI, 1920, there is another article by Tchetwerikov entitled "An experiment on the processing of mass data by means of sampling methods". Vol. VII, 1920, contains a study on "Cattle breeding according to the results of 10 per cent. sample census of agricultural holdings in 1919". In Vol. VIII, 1921, there is an article on "The origin of sampling methods and the first results obtained on the basis of their application in Russia", by A. Gurev. This article is important because it shows that the ideas popularized by Kiaer were used in Russian theory and practice in the second half of the nineteenth century, i.e. much earlier than when Kiaer started his work in this field. Gurev presents many examples to illustrate this statement. In the same volume A. Gozulov reviews the "Results of the sample survey of children aged 0–18 years taken in Rostov and Nakhichevani". Vol. IX contains a report on the results of the 10 per cent. tabulation of the federal agricultural census in the Province of Pskov.

From the above it is clear that before 1920 not only were theoretical articles on sampling methods available for study but practical applications were being made. It is also evident that the use of such methods was not confined to the central authority; they were also employed in different surveys and tabulation programmes in the provincial statistical offices. Already, by that time, sampling had obviously become popular.

In Vol. X, 1922, pp. 179–180, reference is made to the preparation of a pamphlet *Review of the Data on the Tenure of Land*, on the basis of a sample taken in each republic. On p. 181 the same problem is discussed in connection with a sample from agricultural holdings, drawn from other agricultural statistics, data being provided on the allocation of the sample among counties, communes and individual households. On p. 33 of the Appendix, A. I. Krashcheva deals with a sample census of agricultural holdings in 1919 and some information is given about the sample tabulation of the 1920 census. These data form part of a paper which, when read before the federal statistical conference, was followed by an animated discussion regarding the size of the sample needed for the same tabulation. On p. 76 information is given concerning research into the needs of the population for different commodities, the study having been planned as a sample survey.

In Vol. XI, 1922, pp. 151–2, 162, there is another note on the use of sampling methods in agricultural statistics.

Vol. XII, pp. 181–188, mentions, among others, a paper read at a meeting of statisticians in October, 1922, by I. A. Poplavsky, on "The use of sampling methods in transport statistics" in which he suggested that data be collected by means of registration of traffic at certain points of the transport routes. On p. 191 it is stated that the Methodological Division of the Central Statistical Administration has been made responsible for the study of the precision of the results obtained by "mechanical" (i.e. systematic) selection of the sample, and of the question whether the units should be households or clusters of households. In the Appendix, p. 29, A. I. Krashcheva's name appears again. She was probably the first woman to take a deep interest in the theory and practice of sampling methods and is often mentioned in *Vestnik* in connection with them. On this occasion, she had read to the All-Russia Statisticians' Conference, in November, 1922, a report on "Methods of acreage statistics", in which she spoke on the application of sampling methods and drew attention to some attendant theoretical problems requiring study. A. I. Kogen also spoke to that meeting on sampling methods. A resolution (p. 31–32) was passed suggesting the need for further study of a number of problems in the field of sampling methods. The large number reported as speaking on the problem shows in itself how popular such methods then were in Russia.

Vol. XIII gives, on p. 189, an account of a report made to the same meeting by N. L. Dubenetsky on current agricultural statistics advocating the application of sampling methods in that field.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A resolution is also recorded to the effect that "the Methodological Division, in collaboration with the Division of Current Agricultural Statistics, be directed to study the question of the percentage of households that should be drawn into the sample from the whole mass of households . . . and the question of how to select the sample". There is also a reference, on p. 191, to a decision of the Central Statistical Administration whereby the Methodological Division was directed to undertake some experiments in order to clarify the problem of cluster sampling. Some fresh data on sampling methods are to be found on pp. 208–9, 298 and again on p. 300, where a review of Lebedev's report is presented. The title of this report is "The problem of the methodology of sample surveys of agricultural holdings".

In Vol. XVI, A. Krashcheva, on pp. 88–89 again deals with the application of sampling methods, this time in the study of the dynamics of peasant households. The possibility of using such methods in connection with some other problems of agricultural statistics is discussed on pp. 172–175, 178–179 and 184, and again in Vol. XVII, pp. 165–168 and 263–267.

One more item remains to be mentioned, the book *Basic Theory of Sampling Methods*, by A. G. Kowalsky, published in 1924 in Saratov and reviewed in Vol. XX, 1925. I have had no opportunity to see the book itself; from the review it appears to have been very similar to modern text-books on the subject. The author declares that he has confined himself to probability samples because only that type of sample makes it possible to develop an objective and scientific theory of surveys. Another feature of the book is its attempt to give a full mathematical treatment of the topic, and in connection with this many formulae and mathematical derivations are provided. The reviewer also states that some new and unknown theories are presented. One of these (according to the review) is the optimum allocation of the sample among different strata in a stratified population. This theory was introduced into the statistics of western countries by J. Neyman in 1934. Later it was found that Tchuprow had developed the principles of the theory in *Metron* in 1923. Apparently, Kowalsky developed the same theory independently of Tchuprow. He had a formula showing that the best allocation of the total size of the sample is not attained when distribution among the strata is proportional to the number of their units, but, instead, when the stratum sample sizes are proportional to the stratum numbers of units and "the degree of the variation of the phenomenon under study in a given stratum" (p. 260). The reviewer adds that the book includes a new mathematical treatment of the systematically selected samples.

The above exposition is sufficient to prove that Russian statistics should not be neglected in a historical review of the development of the theory and practice of sampling methods. The same holds good for some other fields in which Russian statistics probably then led the rest of the world.

In subsequent years political considerations became an increasingly pronounced factor in the development of Russia's statistics. This brought about the gradual disappearance of the use of theory in the practical activity of the statistical administration. In the late thirties *Vestnik Statistiki* began to close its pages to papers in which statistical problems were dealt with mathematically. At the end of that period they disappeared completely and have not appeared since. The result of this new trend was that statisticians abandoned practice to continue their work at the universities and other scientific institutions where they pursued statistics under the name of some other subject. Officially, Romanovsky, Kolmogorov, Smirnov and many others are mathematicians divorced from statistics. A very interesting example is E. Slutsky, who enjoys world wide renown as one of the forerunners of econometrics. He gave up statistics to embark on a new career in astronomy. Some distinguished names such as Tchetwerikov, Obukhov and a number of others appeared in meteorology. Many younger statisticians are working now as engineers, technicians, and so on. According to official views, statistics became an instrument for planning the national economy. Consequently, its basis is the Marx-Lenin political economy; it represents a social science or, in other words, a class science. The law of large numbers, the idea of random deviations, and everything else belonging to the mathematical theory of statistics were swept away as the constituent elements of the false universal theory of statistical science (see Pisarev, Methodological questions of Soviet Statistics, *Planovoe Hoziaistvo*, 1940; *Vestnik Statistiki*, 1950, which contains a report on the famous conference on methodology, held at the Central Statistical Administration). Nevertheless, the relatively long tradition of sampling methods made it difficult to eliminate them. Thus, S. V. Sholtz, in his text book on agricultural statistics, *Kurs selskohoziaistvenoi statistiki*, Moscow, 1945, claims that such methods represent the best and most objective way of obtaining data about yields. However, sampling methods are now being discussed without any formulae.

It is to be hoped that some day one of the living Russian statisticians who was scientifically active during and after the First World War may write a historical review of Russian statistics, and thus of the development of the theory and application of sampling methods in that country.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REVIEWS OF STATISTICAL AND ECONOMIC BOOKS

CONTENTS

	PAGE
1.—*Sprowls* (*R. Clay*). Elementary Statistics	339
2.—*Vajda* (*S.*). The Theory of Games and Linear Programming	340
3.—*Federer* (*W. T.*). Experimental Design, Theory and Application	340
4.—*Dornbusch* (*S. M.*) & *Schmidt* (*C. F.*). A Primer of Social Statistics	341
5.—*Tippett* (*L. H. C.*). Statistics. 2nd. ed.	342
6.—*Nielsen* (*K. L.*) Methods in Numerical Analysis	343
7.—National Bureau of Standards. Table of Hyperbolic Sines and Cosines $x = 2$ to $x = 10$	344
8.—National Bureau of Standards. Table of the Descending Exponential $x = 2{\cdot}5$ to $x = 10$	344
9.—Harvard University Computation Laboratory. Tables of the Cumulate Binomial Probability Distribution	344
10.—*Fieller* (*E. C.*), *Lewis* (*T.*) & *Pearson* (*E. S.*). Correlated Random Normal Deviates	345
11.—*Neville* (*E. H.*). Rectangular Polar Conversion Tables	346
12.—Ministry of Labour and National Service.—Report on Proposals for a New Index of Retail Prices	346
13.—*Hoffman* (*W.*). British Industry, 1700–1950	347
14.—*Malzburg* (*B.*) & *Lee* (*E. S.*). Migration and Mental Disease	348
15.—London Transport Executive. Health in Industry	349
16.—Economic Commission for Europe. The European Housing Situation	349
17.—*Woytinski* (*W. S.*) & (*E. S.*). World Commerce and Governments	350
18.—*Davidson* (*A. R.*). The History of the Faculty of Actuaries in Scotland 1856–1956	351

1.—*Elementary Statistics.* By R. Clay Sprowls. New York and London: McGraw Hill Book Company Inc., 1955. xiii, 392 pp. 9″. 41*s.* 6*d.*

This elementary statistical text book is written from the point of view that the statistical method is a decision procedure. "Statistics is the science of making decisions based upon incomplete knowledge". Leaving aside the question whether statistics is a science, Sir Ronald Fisher has recently expounded in this *Journal* the view that an important function of statistics is rather the evaluation of the nature and extent of the uncertainty with which decision is encumbered. There is a crucial difference between this evaluation and the making of decisions, and it is unfortunate that the combination of these two activities known as an acceptance procedure should have led to the view that the making of decisions is the business of statistics. The author's view of statistics has also coloured the selection of topics included in the book.

The topics discussed are not developed, however, in terms of statistical decision functions. Such a treatment would be almost impossible for the class of reader the author has in mind. No knowledge of calculus is assumed and consequently very little mathematics is involved. Indeed Professor Sprowls has written this book primarily "not for those who intend to become practising statisticians but for ordinary citizens who will be exposed to statistical statements, arguments and reports". The author's task is therefore exceedingly difficult since the use of mathematical exercises to drive home the statistical concepts is not possible: most of the numerous exercises given are questions on the preceding text involving little thought provided that the text has been understood.

In order, presumably, to stimulate interest and overcome the difficulty of explaining the nature of probability to the non-mathematical reader, Professor Sprowls introduces the subject by means of a horse race problem and draws an analogy with a bead-box experiment. The results of these experiments are quoted in full, but the statistician who has learnt his elements of probability theory with the help of problems on permutations and combinations will probably feel that the bead-box experiments tell only part of the story. The frequency function of the binomial distribution is for example quoted but not derived: had the derivation been included the point that the bead-box experiments actually generate hypergeometric and not binomial samples could have been emphasized and some possibility of confusion avoided: the distinction between sampling with and without replacement occurs only much later in the book. The theory of testing statistical hypotheses follows immediately; the idea that a statistical hypothesis can be rejected but not accepted is not introduced and only the problem of testing a hypothesis against an alternative is discussed. Little is said on the choice of statistics to give the most powerful test; probably this notion could have been more naturally introduced if statistical estimation were discussed before, rather than after, the testing of statistical hypotheses. Further chapters mention descriptive statistics, confidence intervals, and correlation and regression but the ideas of partial correlation

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and regression are not introduced. Succeeding chapters discuss time series and index numbers and there is a final chapter devoted to "further problems"; the concepts involved in these chapters are probably the simplest in the book, but presumably they are considered last as they conform less readily to the author's view of statistics as subject matter for the making of decisions. The section on index numbers contains no mention of Laspeyres or Paasche and the use of index numbers for welfare comparisons is not discussed.

The attempt to explain statistical concepts in simple terms has entailed a rather loose style of writing which at times has led to some misleading statements: "An estimate which is sometimes both unbiased and efficient is the maximum likelihood estimate"; "The least squares method is preferable to the more subjective graphic method only because it leads to a unique set of coefficients"; "A maximum likelihood estimate of an unknown parameter is that value computed from the sample which, if it were the parameter value, would maximize the probability of occurrence of the observed sample"; "A high dividend/earnings ratio was accompanied by a high price/earnings ratio indicating that companies which pay out larger shares of their earnings in dividends have stocks which are high-priced relative to their earnings. This phenomenon may be true, but it is an unwarranted conclusion from the regression *because both ratios have the same denominators*". Again statistical dependence is described as statistical non-independence. Such errors will be spotted immediately by the experienced statistician but might easily cause confusion or misunderstanding to student readers.

Professor Sprowls has emphasized the Neyman-Pearson theory of testing statistical hypotheses at the expense of distribution theory, statistical estimation, and correlation and regression. This treatment accords with his view of statistics but is unfortunate from the points of view of expository convenience and of the historical development of the subject; the discussion of index numbers and the lack of emphasis on regression methods render the book particularly unsuitable for economics students and the emphasis on decision making may well give a false impression of statistical problems to the general reader. The book is on the whole well turned out but there is a distinct asymmetry in some of the normal curves which could well be corrected in a second edition.

F. E. A. BRIGGS.

2.—*The Theory of Games and Linear Programming.* By S. Vajda. London, Methuen (New York, John Wiley), 1956. vi, 106 pp. 6½″. 8*s.* 6*d.*

The reader who is attracted to this book by the first half of its title may find himself rather disappointed. No account is given, for example, of games in extensive form, or of continuous games, or of n-person games ($n \geqslant 3$). In fact, the only games treated are finite rectangular two-person zero-sum games—there is a directory giving the solutions for nine such games, whose matrices are of order 2×2, 2×3, and 3×3. Two proofs of the Minimax Theorem are given, one following von Neumann and Morgenstern, and the other depending on the theory of Linear Programming.

However, the book provides a very welcome elementary introduction to the theory of Linear Programming. It maintains a high standard of rigour, though requiring no mathematical knowledge beyond some elementary algebra. The emphasis is on practical numerical methods and several detailed computing techniques are presented. Good use is made of graphical presentation of problems. The subscript notation is alarming at first ($z_{us u_{m+k}}$ is common and a_{n-11} has two subscripts, $n - 1$ and 1) but it is quite workable, and the descriptions of the methods are remarkably clear. There is an excellent cross-reference system within the book, and an adequate bibliography.

There are about a dozen misprints, mostly unimportant, but the uninitiated reader may be puzzled by a numerical error on p. 6, 1.23 and by the fact that "column" is written for "row" near the end of the next page. He may also be puzzled by a reference on p. 14 to a non-existent proof in a non-existent paragraph IX.4, and by the fact that the first figure on p. 14 is altogether wrong.

These are minor defects, however, and there is no doubt that in the main the book answers its purpose admirably.

J. G. MAULDON.

3.—*Experimental Design, Theory and Application.* By Walter T. Federer. New York: The Macmillan Co., 1956. xix, 544, 47 pp. 9½″. 77*s.*

This book is primarily a student's text-book, and is based on lecture notes for University courses. It contains sixteen chapters, beginning with three introductory ones on the principles of scientific experimentation, significance tests and transformation of data, and field technique. The next twelve chapters relate to various experimental designs and their appropriate analyses, while the final chapter deals with covariance. An appendix contains problems classified by

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

chapters. The 340 papers which are cited in the list of references cover the available literature extensively and include some published as late as 1954.

The examples in the text are heavily biased towards the biological sciences, and agriculture in particular. Probably it is this bias which leads the author to neglect sequential experimentation, which has found applications chiefly outside biology.

In Chapter I, after some very general remarks on scientific experiment, the author gives a classification of experimental designs, and some guidance on the selection of a design and the validity of the experimental error. He is scrupulous in giving the name of the originator of many of the ideas, prefacing his sentences with "X states that . . ." or "Y describes . . ." but this method of writing has its disadvantages. It tends to produce a rather characterless recipe-book of statistics, and this tendency is accentuated when the author does not attempt any evaluation of the work described. Thus he quotes without comment Fisher's expression for the relative efficiency of two designs with errors based on different numbers of degrees of freedom, without mentioning that it is based on a fiducial argument not everywhere accepted (and of a type which the author nowhere else uses), or that other intuitively reasonable measures give rather different answers.

Chapter II contains a discussion of the various tests that have been proposed for assessing the significance of a set of means when no natural partition of contrasts is given by the design of the experiment. A considerable amount of work has been done on this subject recently, and it is useful to have it summarized in one text. It is a pity, though, that the main point of Duncan's multiple range tests, the geometrical progression of protection-levels, is not mentioned.

Chapter III is entitled "Plot or pen technique", and discusses among other things, plot size and shape, the number of replications in an experiment, and sampling of experimental units.

The remainder of the chapters, except the last, contain accounts of various experimental designs and their analysis, beginning with the completely randomized design, proceeding through randomized blocks, latin squares, factorial experiments, and incomplete block designs to change-over designs, with a final chapter on "some additional designs". These latter include, for the first time in any book, the "super magic latin square", an extremely systematic randomized design for the adventurous experimenter.

The general treatment of the material in these chapters is extremely thorough, perhaps at times unnecessarily so. For example, after the student has been shown the form of the analysis for a singly split-plot and a doubly split-plot experiment, he should hardly need to have the trebly split-plot experiment demonstrated to him. If he does need to, he should not be studying statistics. The arrangement of the material in these chapters unfortunately places the details of the analysis of a design before the description of the statistical model involved, when it is this model which alone gives sense to the analysis. The models considered are the additive normal ones in general, and very little is said about the randomization model, and its similarities to and differences from the normal one.

The lattice designs receive an extended discussion with copious examples. The author has made a number of contributions to the analysis of these designs, and his chapters on them undoubtedly form a useful source of reference. Rao's general method of analysis for incomplete block designs is illustrated by its application to the simple and cubic lattices.

The discussion of change-over designs in Chapter XIV is useful because much of the work described is recent, and has not so far appeared in a book. However, the succeeding discussion on long-term experiments is rather too condensed in relation to their importance in agricultural research.

The final chapter covers covariance as applied to most of the designs covered by the previous chapters. Again, the presentation is very detailed and some of the algebra might have been omitted had the student been given at the beginning the basic theory for the fitting of an additional non-orthogonal constant to an orthogonal model.

The standard of printing is very high throughout the book and no errors have been noticed. All tables are reproduced in a "typewriter" script, but are clearly legible none the less. Its price seems likely to put it outside the range of the student, to whom it would probably be of most use.

J. A. NELDER.

4.—*A Primer of Social Statistics.* By S. M. Dornbusch and C. F. Schmid. New York and London: McGraw Hill, 1955. ix, 251 pp. 9″. 35*s.* 6*d.*

The preface to this book describes it as attempting to deal with three main difficulties in the teaching of social statistics: the inadequate mathematical background of the students, the temptation to use a statistical "cook book" (which is largely the result of the former) and the extent of the field which should be covered.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Taking their solution of the last problem first, the authors deal with tabular and graphic presentation; measures of central value and of variability; sampling distributions of proportions, means and measures of correlation; linear regression and correlation, and contingency. The student is introduced to the use of the binomial, normal, t and χ^2 distributions and Fisher's z transformation of the correlation coefficient is described and tabulated. The chapter on "A taste of things to come" at the end of the book introduces the possibilities of multiple and partial correlation, the Guttman scale, analysis of variance and matching. No mention is made of the Poisson distribution, the analysis of time series (even by fitting a least squares line) or of the F distribution. Apart from simple random sampling, there is little discussion of sampling methods and their possibilities and there is nothing on survey methods. In a book on statistics which is intended for sociologists the latter is a surprising omission. The difference between qualitative and quantitative variables is discussed but little attention is given to the difficulties of establishing satisfactory scales of measurement for qualitative variables. Apart from the sociological nature of the illustrative examples and the longer than usual discussion of contingency, this might well be described as a primer of statistics for those with no mathematical background, whatever their main subject of interest.

Considerable attention is given throughout the chapters on sampling to the principles involved and one chapter is devoted to statistical inference. From the point of view of the students for whom this book is intended the latter might have been better placed after the introduction of a statistical test since the logic is at first difficult to appreciate without a particular application in mind. Although no proofs are given of, for example, the sampling variance of the mean and of least squares regression formulae, where possible numerical illustrations are used and some justification is given.

The chief advantage of the book lies in the way it deals with the first problem. Instead of referring the student to a mathematical text the authors aim at giving him the necessary background as he goes along. For those who have forgotten about logarithms, there is a brief refresher explanation in the discussion of semi-logarithmic charts. A whole chapter is devoted to the basic rules of summation and the use of Σ notation which is always a stumbling block. But the authors are not consistent in their treatment. The introduction of percentile range on page 75 is accompanied by a supplementary explanation on "per cent." whereas "percentage change" was elaborated on page 32 and "per cent." was used without comment on page 16. The latter appears in a description of relative frequency distributions which takes less than half a page whereas five chapters later ratios and proportions take up one and a half pages. The whole chapter "Ratios, proportions and rates" would be more appropriately placed before those on central values and variability. In describing a smoothed frequency curve on page 28 a bell-shaped curve is described as a "normal curve" without explanation for the term which appears in the vocabulary at the head of chapter 13. On page 43 pie charts are described but no mention is made of how to make the sectors proportional to the value of the components they represent. A very long (and possibly tedious) proof is given that the sum of the deviations from the mean is zero (pages 62–4) and that

$$\frac{\Sigma(X-\bar{X})^2}{N}=\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2$$

(pages 82–84) while the proofs of the short method for calculating the mean (pages 107–8) and the standard deviation (pages 110–1) receive scant treatment.

In a number of instances, the desire for simplicity leads to inaccuracies. In particular, the method of calculating the denominator of t on page 165 is incorrect. On page 168 it is stated that

$$\text{“}\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{N}}=\frac{S}{\sqrt{N-1}}\text{”}$$

where σ is the standard deviation of the population and s the standard deviation of the sample.

In spite of these deficiencies the student with little or no knowledge of mathematics who requires to learn some statistics will find parts of this text-book well suited to his needs.

R. J. MAURICE.

5.—*Statistics.* By L. H. C. Tippett. 2nd ed. (The Home University Library.) London: Oxford University Press, 1956. vi, 224 pp. $6\frac{3}{4}''$. 7*s.* 6*d.*

One welcomes the second edition of this book, which is so different in character from the typical text-book on statistical methods. This is the subject as it is understood by the layman and it is developed continuously through many possible pitfalls from collecting and presenting material,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

allowing for chance, to the climax of the penultimate chapter called "Statistics in affairs" which implies use primarily in economic, administrative and industrial fields. The book may be treated as an introduction to the subject, for it will be easily intelligible to the beginner, but to my mind it seems even more suitable for those who have acquired some knowledge of statistical methods and an enthusiasm to apply those methods, perhaps a little indiscriminately. As an introduction the book will stand by itself but will have no clear link with the numerical calculations, simple or complex, and the deductive methods taught which make the subject almost entirely a science and not an art. It seems to imply throughout that mathematics is far from being sufficient to make a statistician: there is a stress on the need for experience in dealing with numerical information whilst other books are left to describe most of the mechanical manipulations. Experience often shows that impressions are untrustworthy; potted information such as is given in year books has its dangers as well as uses and, perhaps more stressed than anything else, is the need to know everything about the way the information was obtained and used, as well as something about the circumstances surrounding the events recorded. It is simply standard procedure to examine the results of a sample to see how far they might be explained as random errors, but the real problem is to obtain a satisfactory sample because bias, unlike random errors, cannot be drowned by size. The person who knows statistical methods is much more likely to benefit from such a text, profusely illustrated by examples of errors which have been made, than a person seeking an introduction to this subject.

The first edition of this book, published in wartime, gives a picture of the position before the war, this second edition is post-war in its outlook, recording the effects of the great development which the war has produced, especially in the use of statistics by the government. This development was pictured in the 1943 edition and the following sentence from that edition, with extra words inserted in 1956 recorded in brackets, shows a moderate book becoming more cautious with time—"Planning (in greater or less degree according to the government in power) is the order of the day, and without statistics planning is inconceivable". The number of pages in the book has been increased by over 20 per cent. but this reflects mainly the departure from full austerity of war-time to a much more pleasant lay-out. The chief extension is in the chapter on statistics in affairs which quotes the white paper on Employment Policy as a source of inspiration of many post-war developments in official statistics but perhaps the emphasis of the chapter has been shifted slightly towards industrial uses with mention, or greater mention, of operational research, quality control, linear programming and management accounting. An addition to the chapter on sampling discusses consignment sampling and there are up-to-date sets of diagrams and figures in chapters on arranging and presenting material and on special tables and diagrams. It is difficult to keep such illustrations fully up to date and the old classification of exports is used. Caution is apparently also growing in official statistics; in 1936 road accidents had "causes" but by 1952 there were only contributing factors which the police thought to be "associated with accidents".

This is an unassuming book which teaches the reader more than he thinks; I have been shocked to find how many ideas I thought my own must have been taken from the first edition.

H. S. BOOKER.

6.—*Methods in Numerical Analysis*. By K. L. Nielsen. New York and London: The Macmillan Co. 1956. xiii, 382 pp. 8½″. 48*s*. 6*d*.

This book was written to enable graduate engineers in America to obtain some idea of the nature of the more common numerical methods. It uses only very elementary mathematical concepts for its bases and builds up most of the common tools of numerical analysis from first principles. The aim throughout is to present the results in a form suitable for immediate computational work and most of the methods described can be carried out with an ordinary desk calculator.

The first part of the book is concerned with interpolation into tables of functions using the various common interpolation formulae. This is followed by a discussion of numerical differentiation and integration—a table of the formulae derived and the order of differences to which they are accurate would be of very great assistance at this point. Lagrangian interpolation where a polynomial is fitted to the given points is next dealt with followed by the not altogether related topic of Lagrange's undetermined multipliers.

The second part of the book deals with linear equation systems, differential and difference equations, least squares and periodic functions. The amount of detail given varies enormously from topic to topic. For example, the inversion of matrices and ordinary differential equations are both only dealt with in outline and would often require a much fuller exposition for the practical scientist. On the other hand the fitting of polynomials by least squares is dealt with in great

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

detail, using the ordinary normal equations, weighted or unweighted, a method for equally spaced variates due to Nielsen and Goldstein and the usual orthogonal polynomials. In the final chapter the idea of polynomial representation is extended to trigonometric functions such as Fourier series and exponential types of function. A series of nineteen tables of such constants as those required for Bessel's and Everett's interpolation formulae or for orthogonal polynomials conclude the volume.

As a book designed to accompany a course of lectures few flaws could be found but for a student working alone there seems to be a slight lack of help as to when one or another of the many formulae given is the best to be applied. The examples are mainly straightforward applications of the bookwork and although there are plenty of them it does seem to the reviewer that perhaps too much assistance is given with them. Nevertheless the book is a useful reference work and would be valuable to anyone attending a course on numerical methods. P. G. MOORE.

7.—*Table of Hyperbolic Sines and Cosines* x = 2 *to* x = 10. National Bureau of Standards, Washington, 1955. v, 81 pp. $10\frac{1}{4}''$. 55*c*. (Appl. Math. Ser. 45.)

This table is one in the series of tables of basic mathematical functions in preparation by the Applied Mathematics Division of the National Bureau of Standards.

It gives sinh x and cosh x for $x = 2(\cdot001)10 : 9S$ and supplements previous tables over that range. Linear interpolation will give 6 significant figures and full accuracy can be obtained from the expansions for sinh $(x + p)$ and cosh $(x + p)$ using p and $1 + \frac{1}{2}p^2$ as approximations for sinh p and cosh p respectively.

The tables were produced from these expansions with exact expressions for sinh p and cosh p ($p = \cdot001$) and x at the grid points $2(\cdot05)9\cdot95$. Fifty steps from each grid point produced values which were checked with the next grid value. It is claimed that the tabulated values are not in error by more than one unit in the last place recorded.

Beyond the range both functions behave like $\frac{1}{2}e^x$ to this number of significant figures.

The range $0(\cdot0001)2$ to 9 d.p. has already been published as *Applied Mathematics Series, No.* 36. K. D. TOCHER.

8.—*Table of the Descending Exponential* x = 2·5 *to* x = 10. National Bureau of Standards, Washington, 1955. v, 76 pp. $10\frac{1}{4}''$. 50*c*. (*Appl. Math. Ser.* 46.)

The *Mathematical Tables Project* table of e^x published in 1939 and now available as *Applied Mathematics Series No.* 14 covered the negative range $0(\cdot0001)2\cdot5$ 1(1)100 to 20 d.p.

The present table of e^{-x} covers $2\cdot5(\cdot001)10$ to the same number of places, and will give 6 figure accuracy by linear interpolation. Full accuracy can be obtained by using the addition theorem and value of e^{-x}, $x = 0(\cdot0001)\cdot001$ given in *Applied Mathematics, Series No.* 14.

The table was constructed from the addition table and checked by independent calculation at grid points $2\cdot5(0\cdot1)10$. K. D. TOCHER.

9.—*Tables of the Cumulate Binomial Probability Distribution.* Harvard University Computation Laboratory. Cambridge, Mass.: Harvard University Press, 1955 (London: Cumberlege). lxi, 503 pp. $10\frac{1}{2}''$. 63*s*. (*Annals of the Computation Laboratory* 35.)

This represents the most ambitious tabulation of this function available to date. The specification of the tables is as follows

Function tabulated $E(n, r, p) = \sum_{x=r}^{n} {}^nC_x p^x(-p)^{n-x}$ to 5 decimal places.

$$r = 0(1)n,\ n = 1(1)50(2)\ 100(10)\ 200(20)\ 500(50)\ 1{,}000.$$

$$p\ \cdot0(\cdot01)\cdot50,\ p = \frac{1}{16}\left(\frac{1}{16}\right)\frac{1}{2} \text{ and } p = \frac{1}{12}\left(\frac{1}{12}\right)\frac{1}{2}.$$

The tables are arranged in 6 tables covering the range

(i) $p < \frac{1}{12}$, (ii) $p\ \frac{1}{12} < p \leqslant \frac{1}{6}$, (iii) $\frac{1}{6} < p \leqslant \frac{1}{4}$

(iv) $\frac{1}{4} < p \leqslant \frac{1}{3}$, (v) $\frac{1}{3} < p \leqslant \frac{5}{12}$, (vi) $\frac{5}{12} < p \leqslant \frac{1}{2}$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The multiples of $\frac{1}{16}$ and $\frac{1}{12}$ are inserted into the tables in the appropriate numerical positions and care will be required in p-wise interpolation to note the uneven intervals in p.

The introduction is in four parts of which the first gives the basic properties of the function, derives the recurrence relations on which the tables were constructed and quotes the large n approximations for the function and the bounds on the errors of these approximations.

The second part gives an outline of the methods used to prepare the tables on the Harvard Mark IV calculator which was carried to 12 decimal places. The round-off was accomplished automatically by biasing the initial value in the linear difference relation used to generate the tabular values and each run checked by the formula

$$\sum_{r=0}^{n} E(n, r, p) = 1 + np.$$

The third part gives a complete discussion of the interpolation procedures recommended. For p-wise interpolation the integral form

$$E(n, r, p) = r^n C_r[\int_0^{p_0} x^{r-1}(1 - x)^{n-r}\, dx + \int_{p_0}^{p} x^{r-1}(1 - x)^{n-r}\, dx]$$

is used. Here p_0 is a tabulated value and the second term is evaluated by the three-point Gaussian quadrature formula which can easily be reduced to a simple logarithmic calculation.

Interpolation r-wise is unnecessary and the conventional Bessel formula can be used for n-wise interpolation. Bivariate interpolation should be performed in two steps—n-wise followed by p-wise interpolation.

The last section contains a full account by Mosteller of the application of this function in statistics but does not seem to contain any new material. K. D. TOCHER.

10.—*Correlated Random Normal Deviates.* By E. C. Fieller, T. Lewis and E. S. Pearson. Cambridge, University Press, 1955. xv, 60 pp. $9\frac{1}{4}''$. 10*s*. 6*d*. (*Tracts for Computers*, xxvi.)

These tables have been formed from Wold's well-known random normal deviate tables and give 3,000 sets of 10 deviates $x_0, x_1 \ldots x_9$ such that the correlation between x_0 and x_r is $r/10$.

This constitutes a selection of 25,000 sets of such groups which were constructed by the use of punched cards. The complete set is available on request from the National Physical Laboratory either in punched card form or as a tabulated listing.

The method of production follows a suggestion of Wold. From an original pack of cards containing the Wold number x_i a pack of cards with $\frac{r}{10} x_i$ $r = 1, 2, \ldots 9$ in nine fields was produced. A second pack with

$$\sqrt{\left(1 - \frac{r^2}{100}\right)}\, x_i$$

was also produced (this was distributed over two cards for each x_i) and random numbers were punched on other fields of his second pack. By sorting these random numbers into order new random arrangements independent of the first were obtained and then the sums

$$\frac{r}{10} x_i + \sqrt{\left(1 - \frac{r^2}{100}\right)}\, z_i$$

(z_i representing the rearranged x_i) formed. For each value of r this gave one set of 10 correlated deviates.

These deviates were all subjected to tests for randomness. Those chosen were the sum and sum square test as applied to the original table. In addition, tests of adequacy of the correlation were made by calculating the sample correlations for each set and comparing the sample distribution with the true distribution over samples of size 50, 500, 5,000 and 25,000. This was repeated with sub-sample means. A sign pair change test described in the introduction was also made. The results of all these tests are included in the introduction.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The pairs x_r, x_t have a correlation $\frac{rt}{100}$ and hence the range of possible values of the correlation coefficient available is adequately but irregularly covered.

Although the days of large scale hand sampling experiments are now probably past, there will still be a need for small scale pilot schemes and these tables will form a useful adjunct for the worker in that field. K. D. TOCHER.

11.—*Rectangular-Polar Conversion Tables.* Royal Society Mathematical Tables, 2. Designed and compiled by E. H. Neville. Cambridge University Press, 1956. xxxii, 109 pp. 11″. £1 10*s*.

These are specialist tables prepared for the use of table computers. The aim has been to prepare tables of

$$\theta = \tan^{-1}\frac{y}{x} \text{ and } r(x, y) = \sqrt{(x^2 + y^2)} \text{ and } \log_e r$$

for a set of pivotal integer values of x and y up to 105. They constitute the first major two-entry table project of this kind. Previous tables have reduced the problem to a one-entry table by giving $\tan^{-1} t$ and $u = \sqrt{(1 + t^2)}$ from which the necessary values can be obtained by putting $t = \frac{y}{x}$ and $r = xu$.

In addition to θ in radians, it is also given in degrees and this was calculated independently giving a check on the accuracy of the work.

The relations

$$\tan^{-1}\frac{x}{y} = \frac{\pi}{2} - \tan^{-1}\frac{y}{x} \text{ and } r(x, y) = r(y, x)$$

reduce the necessary tabulation to the triangular range $0 \leqslant y \leqslant x \leqslant 105$. Each page of the main table is devoted to a value m say and tabulates r and θ in degrees for $0 \leqslant y \leqslant x = m$ and also $\log_e r$ and θ in radians for $y = m \leqslant x$. In this way each page contains the same number of entries and the use of the tables is eased by comparison with the usual folded back technique of presenting triangular ranges.

The arithmetic technique used in developing the tables consisted in forming the Farey Series F_{105}. The angles corresponding to the tangents have first differences whose tangents are the reciprocals of integers (using Haros's result concerning Farey series). Using the addition theorem again, a set of tangents whose angles are the second differences can be formed and these are small enough to be evaluated by the Gregory series with very few terms. These angles were expressed in both radians and degrees and the resulting angles built up by repeated summation to form the two θ tables. By eliminating duplication of values in the two difference angle sets the amount of work was considerably reduced below that required to evaluate the 3,375 terms of F_{105}. The calculations were performed to 20 decimal places of which 15 are given in the table.

The introduction gives a thorough account of the methods used to ensure the complete accuracy of the table and this makes fascinating reading. Professor Neville and his co-workers have produced an outstanding example of the table-maker's art. K. D. TOCHER.

12.—*Report on Proposals for a New Index of Retail Prices.* Ministry of Labour & National Service, Cost of Living Advisory Committee. H.M.S.O., March, 1956. Cmd. 9710, iv, 24 pp. 9¾″. 1*s* 6*d*.

Statistical practice in the field of index-numbers is now making rapid strides forward and will soon have caught up with theory all along the line. This is a very welcome, if somewhat belated, development; and nowhere is it more welcome than in connection with the index of retail prices—earlier and more popularly known as the "Cost of Living Index".

The first Cost of Living Index was started during the first world war and was designed to measure the changes in the cost of buying the basket of goods bought by a sample of working-class families in 1904. For various reasons, this basket remained unaltered for nearly half a century, during which time working-class spending habits had undergone substantial changes. Even when the index was revised in 1947 it was not possible to give it an up-to-date weighting pattern: the 1904 basket was replaced by a 1937 basket. Some further revision of weights was made in 1953, on the basis of estimates of changes in total national consumption habits between 1937 and 1950, but it was clear by that time that a new inquiry into the pattern of household expenditure was

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

necessary. This inquiry was made in 1953 and the present report gives us some of its results together with the Cost of Living Advisory Committee's recommendations on the appropriate procedure for constructing a new Index of Retail Prices. These recommendations have been accepted and are now in operation.

The principal difference between the 1953 household budget inquiry and previous ones is that it covered all households irrespective of income or social class. The Cost of Living Advisory Committee were therefore free to choose what part of the results should be used to provide the weights for the new Index. They might, for example, have followed precedent and included only working-class families or they might have gone the whole hog and decided to cover the whole population. They have not given a full account of the considerations they had in mind in making their decision, but we know from a previous report that they were strongly opposed to publishing more than one index and that they were generally inclined to cover as wide a section of the population as could plausibly be covered by a single index. Granted that it is politically advisable to have only one official cost of living index, it is clear that the wider its scope the better. From a statistical point of view, however, this decision can only be justified if the pattern of expenditure does not vary too much between the different economic and social classes covered by the new index. We shall not be able to make a judgment about this until the detailed results of the 1953 Household Budget Inquiry are published.

The Committee decided to exclude only two types of household: those in which the head had a gross income of £20 a week or more, and those in which three-quarters of the total household income was derived from an old age pension or similar payment. The intermediate classes represent nearly 90 per cent. of households and include virtually all manual workers, small and medium salary-earners and self-employed persons earning less than £1,000 a year. It is a striking commentary on the times that it should be thought possible to represent the conditions facing such diverse classes of the community by a single indicator.

The weights of the new Index are based on the quantities consumed in 1953 expressed at the prices of January 1956. The effect of introducing the new quantities has been to reduce the weight of Food and of Fuel and Light in the total index and to increase the weight of other items. The greatest absolute increase is in Transport and Vehicles, most of which is due to the growing importance of the motor-car; but there are appreciable increases in the weights of many other groups. Notable exceptions are Alcoholic Drink and Tobacco. The new weighting system, in fact, focuses attention on trends in our pattern of life of which we are usually but dimly aware. Less and less of our income is going on necessities and more and more on luxuries; and this proves, if nothing else does, that the general standard of life is rising.

Apart from the re-weighting of the Index, the most important recommendations of the Committee concern the extension of the number of separate price quotations to be collected, particularly through the inclusion of the prices of numerous items not previously covered at all. Notable examples of these newly covered goods are rump steak, processed cheese, real coffee, dog biscuits, television sets, nylon stockings, second-hand cars and lipstick. The new Index will be a modern, technically improved, reasonably reliable product and it should last for a fair number of years without more than occasional running repairs. Let us hope that it does not go ahead as fast as its predecessor.

H. F. LYDALL.

13.—British Industry, 1700–1950. By W. Hoffman; translated by W. H. Chaloner and A. O. Henderson. Oxford: Blackwell, 1955. xxiii, 338 pp. 2 fol. tabs. 9″. 35*s*.

Professor Hoffmann's Index of Industrial Production for the United Kingdom since 1700 is well-known, and this book, presenting a full description of the Index, will be welcomed. The original work, in German, was published in 1940; Professor Hoffmann has added a section on production movements for 1935 to 1950, and the translators have added some notes on the statistics and sources. Besides the description of the index the greater part of the book is an analysis of the development of production: the trends, cycles, and correlations shown in aggregate and by the separate industries.

The description of the index is well done. The weights and all the series used as indicators are tabulated in full. The sources of the series are given and the reliability of the series is discussed in the light of other information. Throughout the text there is an ample supply of references.

How reliable is the index? About the early years it seems right to be suspicious; the indicators are very few and the weighting is consequently unbalanced. Wool gets 70 per cent. of the weight in 1740. From about 1850 the information is markedly better and it begins to be possible to check the results. Multiplied by price series the index does quite well when compared to the available national income series; the agreement is not very good, but considering the material even rough

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

agreement is surprising. The index also compares quite well with the other indexes of production available since 1924.

The analytical part of the book, on the other hand, is rather pedestrian and the techniques are not convincing. The long run rates of growth of the individual industries and total production are analysed first, and this is followed by an investigation of cyclical movements, which establishes the existence in most industries of an 18–23 year cycle in the rate of growth. Finally, there is an examination of the long run pattern of industrial development, suggesting that "the typical life-history of an industry was one of steady growth over a long period—characterised by a rising rate of growth of output—followed by a sudden transition to a more or less steady decline in the rate of growth of output. . . . In several industries we were able to identify a third typical phase . . . one of declining output".

Most of the analysis is based on the annual rates of growth shown by re-centred ten-year moving averages. Knowing the awkward habits of moving averages one is left with some suspicion of the twenty-year cycles; the more so when these are found by inspection of graphs and the criterion of what constitutes a peak or trough appears to be rather elastic. Furthermore, while he emphasises the importance of considering each industry according to its own conditions, Professor Hoffman suggests that a general explanation is the replacement cycle for capital goods; but it is then hard to see how the cycle in the production of consumer goods can, as was not infrequently the case, run out of phase with the cycle for producer goods.

Despite the title, Professor Hoffmann gives nearly all his attention to the years before 1914. The years after that are treated as an uninteresting appendix to the nineteenth century, particularly uninteresting because they provide no scope for Professor Hoffmann to demonstrate his twenty-year cycle. For the student of the nineteenth century, however, there is much of interest.

The standard of text and statistics is not always as high as might be desired. Table 6 and diagram 2 would be more readily comprehensible if it were clearly explained that the index excluding building is a simple annual series whereas the other index is a moving average. A "double ten-year moving average" means a ten-year moving average re-centred by using a two-year moving average. On page 89 it is hard to tell how the word "elasticity" is being used. In the comparison of fluctuations in associated industries in table 34, the period quoted as demonstrating a correlation between cotton yarn and piece goods includes the whole of the eighteenth century, during which the two series have a common indicator; there is a similar position with beer and malt.

There is a little introductory discussion of index number problems, but this is mostly elementary and it is disappointing to find nothing on the special problems of constructing indexes over long periods. Here different weightings can make extraordinary differences, and it is not obvious that a chained index is a satisfactory answer or indeed that it has any precise economic meaning at all.

A. A. ADAMS

14.—*Migration and Mental Disease: a Study of first Admissions to Hospitals for Mental Disease, New York, 1939–1941.* By Benjamin Malzburg and Everett M. Lee. New York: Social Science Research Council, 1956. x, 142 pp. 9″. $1.50.

It is generally believed that the incidence of mental disease is higher among migrants than among non-migrants and that the difference is attributable either to selection of poorer risks at the places of origin or to stresses of migration and adjustment at places of destination. The object of this analysis was to study the inter-relationship between migration and mental disease in New York State. The basic data consisted of first admissions to New York State hospitals for mental diseases from July 1st, 1938 to June 30th, 1941, and the first admission to veterans' and licensed hospitals for the last of these three years, the only year for which they are available. The combined first admissions exceeded 46,000 (42,563 whites of whom 15,033 were foreign born and 4,059 non-whites including 536 immigrants) and were classified as dementia praecox, manic depressive psychoses or "all other psychoses" at the time of admission. All comparisons were based on age-standardized rates. The hospital and census data were such that only nativity, colour, sex and age could be controlled, but these include the most important factors governing the rates of mental diseases in various population groups. After examining the adequacy of the material and the problem of diagnosis, consideration is given to differences arising from nationality, place of birth and period of migration.

An examination of admissions in relation to time left no doubt that the risk of first admission to hospital for mental disease was much greater for recent arrivals in New York State than for the combined population of earlier migrants and non-migrants. For both sexes and for residents in each geographic area, the age-standardized rates for white migrants were at least twice those for non-migrants and among non-whites the differentials were even greater. The incidence of mental disease would appear to be high among recent migrants declining with time towards the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

relatively low level of the indigenous population. No firm deduction can be made on this point, however, since a population distribution of migrants by year in New York State is not available. There are indications that in-migration is sometimes followed quickly by out-migration so that an accurate estimate of the population at risk is precarious.

Although it is shown that the rates for mental disease as well as first admissions were greater for migrants than for non-migrants, several qualifications must be borne in mind. Firstly, a higher proportion of the migrants than non-migrants lived in rural areas where it is easier to care for the harmless mentally disordered person at home. Secondly, it is possible that the patients might have been admitted at younger ages in the states from which they came had the same standards of admission applied as in New York State. Thirdly, the same factors that impel migration may also result in mental disease or early stages of mental disease may be accompanied by migration. Lastly, New York State is decidedly atypical and this in itself may make it highly attractive to the less stable individuals. A. BARR

15.—*Health in Industry* . . . London: Butterworth, on behalf of the London Transport Executive. 1956. v, 177 pp. $8\frac{3}{4}''$. 35*s*.

Dr. L. G. Norman and Mr. F. H. Spratling of the London Transport Executive have presented an account of absence from work due to sickness amongst the Executive's staff in the years 1949–52. The groups of staff considered include drivers and male conductors of the central (red) buses, motormen and guards on the Underground, male workshop staff at Chiswick Works, and male and female clerical and technical staff—a total of some 35,000 men and 2,000 women, by far the biggest groups being the drivers and conductors.

Clear definitions are given of all the terms used and the basis of the calculations. Details are given of the medical conditions of recruitment in the different grades, the age at retirement, arrangements for transfer to alternative employment, the hours of work and whether employees are paid during an absence due to sickness—all of which are, of course, important influences on the rate of sickness absence and affect comparisons between groups of workers. Data are given separately for single and married women.

Two main indices are used: the average annual duration of sickness per person—i.e. *how much*, on the average, an employee is away because of sickness in the year, and the annual inception rate (spells), i.e. *how often* he is away. A third index is also used (the quotient of the two just mentioned), the average length of a spell. The bus drivers are singled out for more detailed analysis in terms of frequency distributions. Absences are considered mainly in two groups according to their length—those lasting 1, 2, or 3 days (which do not necessarily require a medical certificate) and those lasting between four days and 6 months.

The basic source of the medical information is the medical certificate given by the employee's general practitioner. All absences of between four days and six months are classified according to the information given on this certificate into 20 "broad diagnostic groups". For each group of employees, and for each "broad diagnostic group", the two main indices referred to above are calculated for each five-year age-group from 20–64, the necessary basic data being given in detail each time. Most of the tables are illustrated in diagrams.

The result is a considerable body of systematically arranged and detailed data on absence due to sickness in the industry, which cannot fail to be a reference book for serious workers in the field of Industrial Medicine for years to come. It will serve as a base-line for comparing the experience of comparable and nearly comparable groups, and is a model of statistical method in this field.

It may be worth pointing out what this book does *not* do, and, of course, never intended to do. There is no discussion of the medical problems arising amongst the employees of the London Transport Executive, and no guidance is given to the reader in the interpretation of the tables and diagrams, which are presented without comment. Study does, of course, reveal interesting facts and suggestions. Examples are: the tendency amongst the operating staff of both buses and underground railways for the younger men to lose more time through short absences than older men; the fact that, in 1951, 60 per cent. of drivers had no absence at all during the year for sickness; the suggestion that middle-aged conductors are more subject to functional nervous disorders than drivers—which is the more trying, the traffic or the passengers? J. A. HEADY.

16.—*The European Housing Situation* (Economic Commission for Europe, Geneva, 1956.) 56 pp. 11″. 3*s*. 6*d*.

This publication consists of monographs for 23 countries giving "a quantitative and qualitative picture of the housing stock in each country and housing requirements and shortages". Almost all European countries (including the U.S.S.R.) are covered. The data are given as far as possible

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

on the basis of the "League of Nations definitions" of dwellings, rooms, etc. (although these definitions are frequently referred to, they are nowhere given. They were drawn up by a Committee of statistical experts in 1938/39). The inclusion in or exclusion from the definition of a room, of the kitchen, is a major, but not the only, difficulty. The data given in the monographs are based on the latest housing censuses together with statistics of building activity since the census.

Attempts at international comparisons are made in a final chapter. For *quantitative* housing standards, calculations have been made showing the number of dwellings in 1954 as a percentage of the increase in population in 1954 and also (in the absence of data on the increase of households), as a percentage of the estimated increase in married couples. The estimate of the latter, however, based as it is on "new marriages, less divorces, deaths and migrations" can only be very approximate in view of the inadequacy of the data for estimating the number of marriages which ceased to exist on account of death, separation and migration and their unreliability as a measure of increase of households. A divorce, or death of a spouse does not necessarily decrease the demand for dwellings. (The table in which these figures are given is unsatisfactory. It is nowhere stated that the figures of births, deaths, natural and actual increase in population, are "per 1000 of the population", nor how many dwellings were built in 1954.) On the basis of these slender estimates the United Kingdom comes out top: the number of dwellings built being 1·7 times the increase in population and 4·2 times the increase in married couples. For *qualititative* housing standards, interesting information is given on "amenities", i.e. the provision of cooking, lighting, heating, bathing and sanitary facilities. In these respects, the Scandinavian countries, Switzerland, the United Kingdom, and Western Germany, are well equipped. In France in 1946, only 36·6 per cent. of the dwellings had inside piped water and 5·9 per cent. a bath, compared with 94·0 per cent. and 62 per cent. respectively in Great Britain (1951). Little information is available on one qualitative aspect of dwellings, viz. size of dwellings and rooms (floor space). Only four countries appear to compile such data and the average floor space varies from 30 sq. metres in Hungary to 60 sq. metres in Belgium. While several countries compile data on the existence of a radio, telephone, refrigerator, no country apparently compiles data on the existence of a garden or cellar (both of which are important in some countries) since no reference is made to these amenities.

While this publication is a useful and concise summary of the available data on the European housing situation in 1954, the international comparisons which are attempted cannot be considered as very satisfactory in view of the differences which exist in the definition of dwelling, household, and room, and in the quality of the dwellings as shown by size (floor space) of dwelling and room, and in the existence or absence of various facilities.

J. W. NIXON.

17.—*World Commerce and Governments.* By W. S. and E. S. Woytinsky. New York: The Twentieth Century Fund, 1955. lii, 907 pp. 10″. $10.

In reading this volume, coming so soon after the companion volume, *World Population and Production*, one's attitude cannot be one of anything other than respectful admiration. Though for a quarter-century citizens of the United States, Professor and Mrs. Woytinsky have retained that superhuman thoroughness, that insatiable appetite for facts, on which German scholarship made its name.

This reviewer invited readers to regard the companion volume as a miniature encyclopaedia (with some reservations about the population statistics); and the same applies to this book. It contains sections on international trade, transport and public finance which will be of direct concern to the statistician. A large part of the book is devoted to recent political history, and to a review of the constitutional and parliamentary structure of the countries of the world; information of this sort should be within the reach of all men holding responsible positions in business or government. Reference books can be relied upon to be up-to-date but jejune, encyclopaedias more abundant but nearly always out-of-date; this book, though falling short of a full encyclopaedia, is a great deal more than a reference book, and comes near to combining the up-to-dateness of the one with the abundance of the other.

Naturally the United States and its economic and political affairs fill a substantial part of the book, though the authors have gone out of their way to make their information world-wide whenever possible. The political and historical chapters are not mere chronicles, but high-grade critical writing. A particularly clear account is given of that event which had such incalculable consequences for mankind, namely the refusal of the United States to join the League of Nations after the First World War.

The comments on taxation are profound and incisive. Readers are advised to turn back to the Old Testament to learn the full subtlety of the story of King Rehoboam, the new advisers whom

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

he took on, and the "whips and scorpions" taxation policy which they recommended. The story is applicable to the modern world. The authors' table of international comparisons of public finance however is spoiled, as are those of the United Nations Statistical Office, by failure to include local government expenditure, which is of much greater importance in some countries than in others. The comparative tables of income tax rates, on the other hand, are clear and valuable.

The chapters on international trade, which open the book, carry a lot of useful information, including an interesting review of nineteenth century material. But the authors supply little material which will enable us to make a judgement on what should be the burning question for present-day administrators, namely how long the present extraordinarily rapid pace of the expansion of world trade in manufactures, at a rate of nearly 7 per cent. per annum, can be expected to last?

The book has some regrettable passages, particularly the chapters dealing with colonies. Just because Germany, before 1914, did little to develop her colonies except as naval bases, and treated them as areas of economic privilege for German importers and exporters, the authors are not entitled to assume that all other countries acted in the same way. British and Dutch colonial rule, though it might have done more, nevertheless did a great deal to improve conditions over wide areas of the world. Being German-Americans the authors have to contend with a double dose of prejudice which prevents their seeing the great benefits brought to the world by Free Trade in the nineteenth century.

COLIN CLARK.

18.—*The History of the Faculty of Actuaries in Scotland* 1856–1956. By A. R. Davidson. Edinburgh: The Faculty of Actuaries, 1956. xviii, 29 pp. $8\frac{1}{2}''$. 20*s*.

Official histories contain certain ingredients: a narrative of events, a discussion of characteristics (even today the author of this one finds it necessary to say what an actuary is), descriptions of personalities, and appendices to carry the detail. The differences lie in the seasoning of the dish and the chef is more important than the ingredients. The histories all fulfil their first purpose, but whether they are of general interest depends on their authors; in this the Faculty has been fortunate.

To many the mere facts of the narrative will be uninteresting: there were periods of activity and of stagnation: men of the past were strangely exercised about issues that seem unimportant today. But there is one event which has a wider than professional interest. The Faculty was born of a disagreement between the English and Scotch members (as they themselves styled themselves) of the Institute of Actuaries. A century-old squabble about professional ethics is unlikely of itself to interest those outside the profession. But the problem of how Scotland and England may best live together is still with us, and looms much larger in Scotland than in England; we have here both a microcosm of the national problem of today, and the story of how, by bridging a serious rift by good sense and goodwill, the lesser problem was solved.

To split a profession is no light thing, and those who do it must feel strongly: to do that they must be fools or men of great character. The early actuaries were not fools, and the author is wonderfully successful in sketching life-size these long dead personalities, and in suggesting the background of Victorian Edinburgh (how different from Victorian London!) in which they lived. Coming to those living or recently dead, the author is more restrained, as is right, but there are some delightful touches whereby he gives a picture of a band of colleagues working for the good of their profession, and those whom it serves.

But the best character sketch of all is the unconscious one the author has given of himself. Here are the gusto his friends know so well; the knowledge of people born of affection and respect for them; the hard criticism applied with a salve; the sudden asides which enliven and enlighten whatever he is about; the jauntiness on the surface and the modesty underneath.

If you like your authors to be urbane and conventional, or if Scotland doesn't interest you, pass this book by. If you enjoy firm character sketches, or want to capture something of the taste of life in Edinburgh, dip into it. If you are interested in actuaries and their work, read it. But read it above all for salty writing reflecting the forceful and generous personality of the author.

M. D. W. ELPHINSTONE.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL NOTES

(1) British Official Statistics

The index of retail prices compiled by the Ministry of Labour and National Service which was 101 (January 17th, 1956 = 100) in March rose to 103 in April and May. The increase in April was due mainly to rises in the prices of potatoes and other vegetables and fruit and to some in rents and rates, these being partly offset by reductions in the average prices of eggs and butter. The detailed figures of the weights used in calculating the index and the indices for the different commodity groups are as follows:—

Date	*Food*	*Alcoholic Drinks*	*Tobacco*	*Housing*	*Fuel and Light*	*Durable Household Goods*	*Clothing and Footwear*	*Transport and Vehicles*	*Miscellaneous*	*Services*	*All Items*
Weights:	350	71	80	87	55	66	106	68	59	58	1,000
March 13th, 1956	102·9	100·0	100·0	100·5	100·0	100·3	100·4	100·3	101·1	101·4	101·3
April 17th ,,	106·0	100·0	100·0	102·9	100·1	101·0	100·5	100·9	101·5	101·8	102·7
May 15th ,,	104·2	100·7	105·3	102·7	98·0	101·2	100·4	101·9	101·6	102·1	102·5

A supplementary calculation shows that, on the basis of June 1947 = 100, the index at May 15th was 157, compared with 158 at April 17th and 155 at March 13th.

The Ministry of Labour index of weekly wage rates which was 162 in March (June 30th, 1947, = 100) rose to 163 in April and 164 in May. The principal classes of workers who received increases in April included building and civil engineering workers, cotton operatives and employees in retail furnishing. In May increases mainly affected the mercantile navy and printing operatives.

The total working population and the numbers in civil employment in the three months ended April, 1956, were as follows:

(*Thousands*)

	Total Working Population			*Numbers in Civil Employment*		
	Males	*Females*	*Total*	*Males*	*Females*	*Total*
February, 1956 .	16,128	7,860	23,988	15,190	7,761	22,951
March ,, .	16,116	7,836	23,952	15,211	7,742	22,953
April ,, .	16,128	7,847	23,975	15,226	7,755	22,981

The number of persons on the unemployment registers of the Employment Exchanges fell by 14,000 in April and by a further 14,000 in May. The total for May represented 1·1 per cent. of the estimated total of employees in Great Britain. The percentage in the different Regions ranged from 0·6 in the North Midlands and 0·7 in London and South-Eastern and the East and West Ridings to 1·9 in Wales and 2·4 in Scotland.

The following is a sex analysis of the figures:

Number of Unemployed Persons on the Registers of Employment Exchanges

Date	*Men and Boys*	*Women and Girls*	*Total*
March 12th, 1956 . . .	175,903	89,619	265,522
April 16th ,, . . .	161,519	89,996	251,515
May 14th, ,, . . .	153,388	84,043	237,431

Of the total of 237,431 in May, 62,210 had been unemployed for not more than two weeks, 52,956 for two to eight weeks, and 92,737 for more than eight weeks, while 29,528 were temporarily stopped.

The number of vacancies filled by the Employment Exchanges in the four weeks ended May 2nd was 237,396 and there were 380,328 unfilled vacancies on the registers.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The numbers of insured persons absent from work owing to sickness, including self-employed as well as employed, were 1,062,600 in March, 912,900 in April and 881,000 in May. The numbers of employed persons absent owing to industrial injuries were 61,800, 57,600 and 59,100.

After an interval of fifteen years the Ministry of Labour has again made an enquiry into the working of profit-sharing and co-partnership schemes. It was found that at the end of 1954 there were 421 schemes operating on a pre-arranged basis covering firms with 611,713 employees, of whom 389,433 were participants. In addition there were 130 schemes covering 132,672 participants of a broadly profit-sharing character.

The *Ministry of Labour Gazette* for June, 1956, contains an age analysis of the employee population of Great Britain in May, 1955. Of the total of 13,960,000 males 194,000 were 70 and over and 568,000 were 65 and over. Amongst 7,500,000 females, 42,000 were 70 and over and 137,000 were 65 and over. Figures relating to married women showed a total of 3,570,000 or nearly 48 per cent. of the total number of female employees. The number had increased by 320,000 in a year. The review also contains regional and industrial analyses of the figures and some estimates of inter-regional migration.

The analysis by the General Register Office of the 1951 Census figures on usual residence and workplace has been published under the title *Census* 1951, *England and Wales, Report on Usual Residence and Workplace* (H.M.S.O. £3 5*s.* 0*d.* net). The census question about usual residence was repeated in 1951 in similar terms to those used in the 1931 Census, when it was first asked. Just over 1,000,000 people in England and Wales stated that their usual residence was outside the borough or the urban or rural district in which they were enumerated: this was 2·3 per cent. of the total enumerated population of 43,750,000 and compares with 2·1 per cent. in 1931.

Visitors from outside England and Wales numbered 107,000 of whom 24,000 were enumerated in the County of London. Some 28,000 came from Scotland, whereas the 1951 Census of Scotland recorded 23,000 visitors from England and Wales. There were also 7,000 from the United States, 6,000 from Australia, and 6,000 from Germany.

The tables also show the extent to which persons enumerated and persons resident differed in various local authority areas. Thus, even as early in the year as April (when the census was taken), the resident population of Bournemouth was over 4 per cent. less than the enumerated population, while those enumerated in Oxford (the university being on vacation) were 8 per cent. less than the resident population (which includes the undergraduates). These instances, however, are two of the largest differences. For, while holiday resorts or towns with large institutions or ports with visiting seamen had some visitors and school and college towns and defence stations had residents temporarily absent, the differences between the number enumerated and those resident were relatively very small in the great majority of areas.

The desirability of choosing a time when differences are very small is of course one of the considerations taken into account in fixing the Census date.

The 1951 census question about the situation of the place of work of the people who are gainfully occupied was similar to that asked at the 1921 Census, the only other census in this country for which such information has been published. In the tables for 1951, however, the information has been related to the area of usual residence, and not to the area of enumeration as in 1921.

The tables show, for every borough, and every urban and rural district, how many persons resident in an area had a workplace in some other area, and vice versa. This enables a comparison to be made between the daytime and night time populations of each area, and shows how in general the industry of most towns depends upon workers who live in the surrounding areas. Partly the workers have tended to move farther from their offices or factories and partly the towns cannot find homes to balance the growing number of jobs available.

In London, the City and six central boroughs had a daytime population of 1,500,000, against a resident population of 500,000, or three times as much, whereas in 1921 the resident population was nearly 750,000, but the daytime population only slightly greater than in 1951, that is, only doubled compared with the resident population. The City itself had a daytime total of 340,000 against a resident total of only 5,000.

In Manchester, at the centre of the South East Lancashire conurbation, the daytime numbers exceeded the night or resident population by 74,000, 22 per cent. of the occupied population enumerated within it.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Other areas outside the conurbations with big differences were Derby, with 33,000 daytime excess, Oxford, nearly 15,000 and Doncaster, nearly 10,000.

The report also points to some of the areas which are important dormitories for the towns which import labour during the day, and gives an impression of the wide range of journeys undertaken. It should thus be of interest both to housing and to transport authorities.

The Ministry of Fuel and Power Statistical Digest, 1955 (H.M.S.O. 1956, £1 7*s.* 6*d.*) is the sixth in the present series. The 164 tables cover the production and consumption of coal, electricity gas, coke, manufactured fuel and petroleum. Three new tables in this issue deserve special mention. For the first time, the Digest contains a table of total primary fuel consumption in terms of coal equivalent (table 5) and, for the post-war years, consumption in terms of coal equivalent is broken down by types of fuel for nine consumer groups (table 6). Table 3 shows the value of fuel imports and exports for each year since 1923; these data have not previously been easily available.

(2) Other Statistics

In 1952 the Organisation for European Economic Co-operation decided to make an enquiry into probable population trends in member countries. In 1953 these countries were invited to make forecasts by sex and age of trends in the population between 1951–71 and the request was accompanied by suggestions as to assumptions and methods designed to achieve comparability. The replies were examined by a group of demographic experts and a summary report was prepared. This has now been published by O.E.E.C. (Paris) under the title of *Demographic Trends in Western Europe* 1951–71.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL AND ECONOMIC ARTICLES IN RECENT PERIODICALS

UNITED KINGDOM—

Accounting Research—

April 1956—Depreciation, cost allocation and investment decisions: *H. R. Anton.* Balance and bottlenecks: *J. A. Scott.* A financial statement computer: *R. H. Gregory* and *T. A. Mangelsdorf, Jr.* Some aspects of post-war company finance: *A. Luboff.*

Annals of Human Genetics—

May 1956—A test for segregation ratios in family data: *C. A. B. Smith.* A contribution to the study of birth weight based on an Italian sample: *M. Fraccaro.* The estimation and significance of the logarithm of a ratio of frequencies: *J. B. S. Haldane.* Hair-colour variation in the United Kingdom: *E. Sunderland.* A note on additional material from three regions: *N. A. Barnicot.*

Biometrika—

June 1956—Studies in the history of probability and statistics. II. The beginnings of a probability calculus: *M. G. Kendall.* On estimating the latent and infectious periods of measles. I. Families with two susceptibles only: *N. T. J. Bailey.* The behaviour of an estimator for a simple birth and death process: *J. H. Darwin.* Examination of a quantum hypothesis based on a single set of data: *S. R. Broadbent.* The number of new species, and the increase in population coverage, when a sample is increased: *I. J. Good* and *G. H. Toulmin.* A sequential test of randomness for events occurring in time or space: *D. J. Bartholomew.* On the moments of the maximum of partial sums of a finite number of independent normal variates: *A. A. Anis.* On the application to statistics of an elementary theorem in probability: *H. A. David.* χ^2 probabilities for large numbers of degrees of freedom: *J. Wishart.* The sampling distribution of a maximum-likelihood estimate: *J. B. S. Haldane* and *S. Maynard Smith.* Tests for randomness of points on a line: *D. E. Barton* and *F. N. David.* Paired comparison designs for testing concordance between judges: *R. C. Bose.* On the distribution of the largest or the smallest root of a matrix in multivariate analysis: *K. C. S. Pillai.* Tests of significance for the latent roots of covariance and correlation matrices: *D. N. Lawley.* The variance of the mean of systematic samples: *R. M. Williams.* Grouping methods in the fitting of polynomials to unequally spaced observations: *P. G. Guest.* On the joint distribution of the circular serial correlation coefficients: *G. S. Watson.* The approximate distribution of serial correlation coefficients: *H. E. Daniels.* Tests of hypotheses in the linear autoresgressive model. II. Null distributions for higher order schemes: non-null distributions: *G. M. Jenkins.* A class of distributions for which the maximum likelihood estimator is unbiased and of minimum variance for all sample sizes: *D. E. Barton.* Further critical values for the two-means problem: *W. H. Trickett, B. L. Welch* and *G. S. James.* An approximation for the symmetric, quadrivariate normal integral: *J. A. McFadden.* Weighted probits allowing for non-zero response in the controls: *M. J. R. Healy.* Treatment variances for experimental designs with serially correlated observations: *J. C. Bucher.* The multivariate distribution of complex normal variables: *R. A. Wooding.* Stationarity conditions for stochastic processes of the autoregressive and moving-average type: *J. Wise.* Some properties of an angular transformation for the correlation coefficient: *B. I. Harley.* Note on the moment-problem for unimodal distributions when one or both terminals are known: *C. L. Mallows.* On inverting a class of patterned matrices: *S. N. Roy* and *A. E. Sarhan.* A note on the risks of error involved in the sequential ratio test: *J. Mehdi.*

British Journal of Preventive and Social Medicine—

April 1956—Mortality in the London Boroughs, 1950–52, with special reference to respiratory disease: *D. Hewitt.* Scottish data on the association between physical activity and severity of paralysis in poliomyelitis: *I. N. Sutherland.* Human biology and human speech: *L. Hogben.* Study of sex, age, and regional differences in the advantage of rural over urban mortality: *W. J. Martin.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cambridge Philosophical Society, Proceedings—

April 1956—The numerical evaluation of a class of integrals: *P. A. P. Moran.* Statistical properties of a moving wave-form: *M. S. Longuet-Higgins.* On clumps formed when convex laminae or bodies are placed at random in two or three dimensions: *C. Mack.*

Economica—

May 1956—Some notes on the estimation of time-forms of reactions in interdependent dynamic systems: *A. W. Phillips.* A study in price policy: *I. F. Pearce.* The abolition of derating: an exercise in differential incidence: *A. Williams.* The origins of the synthetic alkali industry in Britain: *T. C. Barker, R. Dickinson,* and *D. W. F. Hardie.*

Institute of Actuaries, Journal—

Vol. 82, Part I—Mortality of assured lives: a memorandum submitted by the Joint Mortality Investigation Committee. Can we express $\bar{e}_{xxx}$ . . . (*n*) and I/μ_x by the ordinates of frequency curves?: *Sir W. Elderton.* Laws of mortality which satisfy a uniform seniority principle: *T. N. E. Greville.* The geometric logarithm and discrete Pareto forms of series: *P. G. Moore.* The recent trend of mortality in Great Britain: *W. S. Hocking.*

Journal of Agricultural Economics—

June 1956—Productivity and costs in Australia and New Zealand: *Colin Clark.* Beef supplies in the United Kingdom—retrospect and prospect: *R. A. Mackness.* A study of economic costs of production: an analysis of horse and tractor costs: *K. Dexter.* The demand for frozen beef in the United Kingdom, 1880–1940: *R. Duncan.*

Manchester School of Economic and Social Studies—

May 1956—Wages: industry rates, workplace rates and the wage-drift: *H. A. Turner.* Exports and the British trade cycle: *D. C. Corner.* Econometric models and the evidence of time series analysis: *L. R. Klein.*

Oxford Economic Papers (New Series)—

June 1956—A note on Mr. Kaldor's trade cycle model: *J. Black.* Wages in coal-mining since nationalization: *K. J. W. Alexander.* The genesis of the multiplier theory: *A. L. Wright.* The commercial crisis of 1857: *J. R. T. Hughes.*

INDIA—

Calcutta Statistical Association Bulletin—

December 1955—On some aspects of non-regular testing of hypotheses: *K. C. Chanda.* Use of the reciprocal of a normal variate in textile statistics: *J. N. Purohit* and *R. M. Phatarford.* Measures of over-all efficiency of sample multinomial tables: *S. J. Poti.* Analysis of latinized rectangular lattice designs: *P. M. Roy.* Sampling inspection of procedures: *P. K. Banerjee.*

Sankhyā—

April 1956—Biometry: *J. B. S. Haldane.* Is biometry a separate discipline?: *H. Spurway.* A problem in the sequential design of experiments: *R. Bellman.* National Sample Survey Number Seven: couple fertility.

UNION OF SOUTH AFRICA—

South African Journal of Economics—

March 1956—The immigration cycle in Palestine and Israel, 1919–1954: *Hon. E. Samuel.* The accuracy of forecasts in the S. African business opinion surveys: *N. van den Berg* and *C. Verburgh.*

UNITED STATES—

American Statistical Association, Journal—

March 1956—Confession of faith: 1955: *R. J. Watkins.* Statistician and policy maker: a partnership in the making: *W. Z. Hirsch.* A test of the accuracy of a production index:

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

C. F. Carter and *M. Robson*. On simplifications of sampling design through replication with equal probabilities and without stages: *W. Edwards Deming*. Investigating the properties of a sample mean by employing random subsample means: *H. L. Jones*. Some theoretical aspects of the lot plot sampling inspection plan: *L. E. Moses*. The operating characteristic curve for sequential sampling by variables when the producer's and consumer's risks are equal: *N. Garner*. Table of percentage points of Kolmogorov statistics: *L. H. Miller*. Multiple regression with missing observations among the independent variables: *G. L. Edgett*. On linear combinations of several variances: *B. L. Welch*. Tabular analysis of factorial experiments and the use of punch cards: *J. R. Bainbridge*, *A. M. Grant* and *U. Radok*.

American Statistician—

April 1956—Some statistical aspects of a problem in scientific philosophy: *J. E. Freund* and *I. Miller*. An empirical study of an estimate of skewness for the Pearson type III distribution: *J. Caffrey*, *N. C. Perry* and *D. Teichroew*.

Econometrica—

April 1956—A generalization of the von Neumann model of an expanding economy: *J. G. Kemeny*, *O. Morgenstern* and *G. L. Thompson*. Approaches to the bargaining problem before and after the theory of games: a critical discussion of Zeuthen's, Hicks's and Nash's theories: *J. C. Harsanyi*. The determination of subjective characteristic functions in games with misperceived payoff functions: *R. Duncan Luce* and *E. W. Adams*. Identification of a certain stochastic structure: *H. Teicher*. Semi-orders and a theory of utility discrimination: *R. Duncan Luce*. Note on initial conditions in the solution of linear differential equations with constant coefficients: *R. H. Tizard*.

Estadística—

September–December 1955—III. Inter-American Statistical Conference, Quitandinha, Petropolis, Brazil, June 9–23, 1955: summary and resolutions.

Harvard Business Review—

May–June 1956—Mathematics for decision makers: *R. K. Gaumintz* and *O. H. Brownlee*. Guides to inventory policy. III Anticipating future needs: *J. F. Magee*. How to identify promotable executives: *C. W. Randle*.

Industrial Quality Control—

March 1956—Practical linear programming applications: *H. T. Schwan*. Linear programming: *A. W. Tucker*. The use of linear programming to improve the quality of decisions: A. C. Rosander.

May 1956—Integrating operations research into a business: *H. D. Mills*. Management of the quality control function: *A. V. Feigenbaum*. Quality control engineering: *W. J. Masser*. Quality control, industrial engineering and operations research: *W. E. Alberts*. An overall quality assurance plan: *E. G. D. Patterson*. Installation of a quality control program: *J. E. Rutledge*. Studies in textile testing: *N. R. Garner*. The use of continuous sampling in ammunition procurement: *R. L. Storer*. Correlation analysis in batch process control: *J. D. Hinchen*.

Journal of Experimental Education—

March 1956—The inductive-deductive method and the physical science laboratory: *A. M. Lahti*. Empirical comparison of six cessation tests for use in principal components factor analysis: *J. E. Stecklein*. The stability of equated test scores: *B. P. Karon*. Group-study versus lecture-demonstration method in physical science instruction for general education college students: *J. N. Ward*. Analysis-of-variance models and their use in a three-way design without replication: *D. M. Medley*, *H. E. Mitzel* and *A. N. Doi*. The multiple group method of factor analysis and rotation to a simple structure hypothesis: *P. Horst* and *K. W. Schaie*. An analysis of variance method for determining the external and internal consistency of an examination: *W. J, Moonan*. Sampling error due to choice of split in split-half reliability of coefficients: *F. M. Lord*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Journal of Political Economy—

April 1956—The welfare cost of inflationary finance: *M. J. Bailey.* Controlled inflation in Israel, 1949–54: *M. E. Kreinin.* Econometric models of the United States: *K. A. Fox.* "Availability" and other influences on the commodity composition of trade: *I. B. Kravis.*

Management Science—

January 1956—Computers as tools for management: *J. S. Coleman.* Formulating a theory of work measurement: *A. Abruzzi.* Recent advances in linear programming: *G. Dantzig.* An inventory policy for a case of lagged delivery: *M. Beckman* and *R. Muth.* Current status of the industrial use of linear programming: *L. Wheatson Smith, Jr.* Derivation of a linear decision rule for production and employment: *C. G. Holt, F. Modigliani* and *J. F. Muth.* The objective of TIMS (Presidential Address): *M. M. Flood.* Management science—fact or theory?: *C. West Churchman.*

Milbank Memorial Fund Quarterly—

April 1956—Fertility and research on fertility in Japan: *I. B. Taeuber.* Mortality declines and age distribution: *G. J. Stolnitz.*

Operations Research—

April 1956—Of horseless carriages, flying machines, and operations research: a tribute to Frederick William Lanchester (1868–1946): *J. F. McCloskey.* A Monte Carlo model of Lanchester's square law: *I. Driggs.* Multiple factor break-even analysis: the application of operations-research techniques to basic problem of management planning and control: *R. S. Weinberg.* A brief history of the development of Canadian military operational research: *N. W. Morton.* A graphical solution of the transportation problem: *M. L. Vidale.* Optimum ammunition properties for salvos: *J. E. Walsh.* Queueing with special service: *E. Koenigsberg.* Waiting-line theory as a management tool: *W. R. van Voorhis.* The determination of requirements for warehouse dock facilities: *D. H. Schiller* and *M. M. Lavin.*

Psychometrika—

March 1956—The additive constant problem in multidimensional scaling: *S. J. Messick* and *R. P. Abelson.* A rapid non-parametric estimate of multi-judge reliability: *D. S. Cartwright.* A study of speed factors in tests and academic grades: *F. M. Lord.* Optimal test length for maximum differential prediction: *P. Horst.* The regression of gains upon initial scores: *R. F. Garside.* A method of scalogram analysis using summary statistics: *B. F. Green.* Note on Carroll's analytic simple structure: *H. F. Kaiser.*

Quarterly Journal of Economics—

May 1956—International wheat agreements and problems, 1949–56: *H. C. Farnsworth.* A general theory of executive compensation based on statistically tested propositions: *D. R. Roberts.*

Review of Economics and Statistics—

May 1956—The relation between home investment and external balance in the light of British experience, 1945–1955: *R. Nurkse.* Does productivity rise faster than in the United States?: *G. D. A. MacDougall.* Major developments affecting the United States balance of international payments: *W. Lederer.* The statistical foundations of the gross national product: *G. Jaszi.* The effect of aggregate versus per capita income per unit ranking on personal income size distributions: *B. Clyman.*

AUSTRIA—

Statistische Vierteljahresschrift—

Vol. VIII, Part 4—Die Anwendung der Statistik in der Nationalökonomie und der Physik: *S. Sagoroff.* Tabakkonsum und Erkrankungen der Atmungsorgane: *W. Ossadnik.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

DENMARK—

Nationaløkonomisk Tidsskrift—

Vol. 94, Parts 1–2—Aggregate consumption and investment functions for the household sector considered in the light of British experience: *R. Stone* and *D. A. Rowe.* Strukturændringer i Dansk ökonomi: *P. N. Andersen.*

FRANCE—

Journal de la Société Statistique de Paris—

January–March 1956—La critique historique appliquée aux documents statistiques et numériques: *J. Bourdon.* Le plan "à cinq dimensions" pour l'élaboration du questionnaire: *J. Antoine.*

Population—

January–March 1956—La croissance des classes jeunes et le problème de l'emploi: *J. Fourastié.* Le remariage des veufs et des veuves: *R. Pressat.* Evolution démographique en Espagne: *J. Daric.*

Revue de Statistique Appliquée—

Vol. IV, No. 1—Différents aspects du contrôle statistique dans l'industrie automobile: *J. Ronchaud.* Détermination d'un mode de prélèvement pour contrôler la qualité fonderie d'après les résultats de l'usinage en mécanique: *M. Courtillot.* Essai de contrôle par prélèvements sur les presses à emboutir les éléments de carrosserie automobile: *M. Duffour.* Les applications du contrôle statistique aux États-Unis: *R. Rambach* and *R. Cave.* La Recherche Opérationnelle: *E. Ventura.* État de la Recherche Opérationnelle aux États-Unis en 1955: *Y. Fournis.* Étude de la propagation des arrêts aléatoires dans les chaînes de production: *B. Zimmern.*

GERMANY—

Allgemeines Statisches Archiv—

Vol. 40, Part 1—Der Lebenszyklus in Erwerbstätigkeit, Einkommensbildung und Einkommensverwendung: *H. Schmucker.* Über Fehler, Fehlerausgleich und Fehlerfortpflanzung in der Sozialstatistik: *H. Kallmeyer.*

Weltwirtschaftliches Archiv—

Vol. 76, Part 2—Experiment und Berechnung grossen Umfangs in der Wirtschaftswissenschaft: *O. Morgenstern.* Economic development in Central America: *B. F. Hoselitz.* Die Wachstumstendenzen des west-deutschen Arbeitspotentials 1925 bis 1955: *H. Wander.*

ITALY—

Statistica—

January–March 1956—La popolazione nel pensiero di Antonio Genovesi: *P. Fortunati.* La dissomiglianza: *G. Landenna.*

JAPAN—

Memoirs of the Faculty of Science, Kyusu University, Series A (Mathematics)—

March 1956—Fluctuations of sums of independent random variables: *N. Ikeda.* Linear translatable stochastic functional equations and random distributions: *T. Onoyama.*

SPAIN—

Trabajos de Estadistica—

Vol. VI, Part III—Sobre la distribución de los tamanõs de corpúsculos contenidos en un cuerpo a partir de la distributión en sus secciones o proyecciones: *L. A. Santalo.* Un problema de minimo planteado en Arquitectura: *P. Zoroa.* Aportión al estudio de un método rápido de cálculo para la aplicación del método del Simplex al problema del transporte de Hitchcock: *J. de Lasala.* Análisis discriminate de dos muestras de indios venezolanos: *A. Diaz Ungria, A. Camacho* and *S. Rios.* Conceptos operacionales de la inspec-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ción: *E. Blanco.* La teoría de las colas aplicada a un problema de producción industrial: *J. Romani.* El análisis operacional en la propaganda: *P. Zoroa.*

SWITZERLAND—

Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker—

April 1956—Ein sequenz-Test zur Kontrolle von Ausscheidehäufigkeiten: *G. Wünsche.* Über eine allgemeine Methode der Lösung des Zinsfuss-problems für verschiedene Versicherungsformen und die Darstellung der darin auftretenden Momente: *H. Gubler.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

LIST OF ADDITIONS TO THE LIBRARY

Since the issue of Part II, 1956, the Society has received the publications enumerated below.

I.—OFFICIAL PUBLICATIONS

United Kingdom. *Central Statistical Office.* National income statistics: sources and methods. London, H.M.S.O., 1956. viii, 387p. 9½″. 25*s.* (Stud. Off. Statist. 3)

Department of Scientific and Industrial Research. Automation: a report on the technical trends and their impact on management and labour. London, H.M.S.O., 1956. viii, 106p. photos. 9¾″. 6*s.*

General Register Office. Census 1951. Classification of occupations. London, H.M.S.O., 1956. xvi, 317 p. 13¼″. 30*s.* Revision of *Classification of occupations 1950.*

Census 1951. England and Wales. Report on Greater London and five other conurbations. London, H.M.S.O., 1956. cxxxv, 351p. 6 fold. maps. 13¼″. 105*s.*

Census 1951. England and Wales. Report on usual residence and workplace. London, H.M.S.O., 1956. xxxix, 268p. 13¼″. 65*s.*

General practitioners' records: an analysis of the clinical records of some general practices during the period April 1952 to March 1954. London, H.M.S.O., [1956]. vii, 80p. 9½″. 6*s.* 3*d.* (Stud. Med. Popul. Subj. 9, in continuation of Stud. 7)

Matters of life and death. [3rd ed.] London, H.M.S.O., 1956. 28p. 8½″. 1*s.*

Medical Research Council. B.C.G. and vole bacillus vaccines in the prevention of tuberculosis in adolescents: first (progress) report to the Medical Research Council by their Tuberculosis Vaccines Clinical Trials Committee. *Brit. Med. J.* (1956), **1**, p. 413. 15p. 11″.

Ministry of Agriculture, Fisheries and Food. Report of the Committee appointed to review the provincial and local organisation and procedures of the Ministry of Agriculture, Fisheries and Food. London, H.M.S.O., 1956. vi, 107p. map. 9¾″. 4*s.* 6*d.* (Cmd. 9732)

Ministry of Health. Blindness in England, 1951–1954; by Arnold Sorsby. London, H.M.S.O., 1956. vi, 53p. 9½″. 3*s.*

Ministry of Health & Ministry of Housing and Local Government. The bacteriological examination of water supplies. London, H.M.S.O., 1956. viii, 52p. 9¾″. 2*s.* 6*d.* (Rep. Publ. Hlth. & Med. Subj. 71)

Ministry of Housing and Local Government. Differential rent schemes. London, H.M.S.O., 1956. 11p. 9¾″. 6*d.* (Circular 29/56)

Rates and rateable values in England and Wales. Supplement to the 1955–1956 edition: showing rateable values and numbers of hereditaments in the new valuation lists which came into force on 1st April, 1956. London, H.M.S.O., 1956, 55p. 9¾″. 3*s.*

Ministry of Pensions and National Insurance. National Insurance Act, 1946. Report of the National Insurance Advisory Committee . . . on the question of earnings limits for benefits. London, H.M.S.O., 1956. 54p. 9¾″. 2*s.* (Cmd. 9752)

National Physical Laboratory. Wage accounting by electronic computer: report no. 1 of the Inter-Departmental Study Group on the application of computer techniques to clerical work. London, H.M.S.O., 1956. [ii], 57p. 9¾″. 2*s.* 6*d.*

Parliament. The economic implications of full employment. London, H.M.S.O., 1956. 13p. 9¾″. 9*d.* (Cmd. 9725)

Treasury. Organisation and Methods Division. Creating the punched card. [London], Treasury, 1953. i, 48p. 13″.

Australia. *Bureau of Agricultural Economics.* The Australian sheep industry survey, 1954. Queensland. Canberra, 1956. [vi], 26p. map. 9½″.

Austria. *Osterreichisches Statistisches Zentralamt & Österreichisches Institut für Wirtschaftsforschung.* Forschungsstelle zur Aufstellung volkswirtschaftlicher Bilanzen. Der Verbrauch der städtischen Bevölkerung Österreichs: Ergebnisse der Konsumererhebung 1954/55. Wien, 1956. viii, 127p. Blank form and notes in pocket. 12″.

Eire. *Central Statistics Office.* Census of population 1956. Preliminary report. Dublin, 1956. 9 l. 13″.

Finland. *Statistiska Centralbyrån.* Väestötilastoa . . . 1950 population census. Vol. III. Housing. Helsinki, 1956. [ii], 136p. 11¾″. (Official Statistics of Finland VI. C 102)

Väestötilastoa . . . Vital statistics in Finland for the year 1954. Tables. Helsinki, 1956. 64p. 11¾″. (Official Statistics of Finland VI. A 112)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Germany. *Statistisches Bundesamt.* Die berufliche und soziale Gliederung der Bevölkerung der Bundesrepublik Deutschland nach der Zählung vom 13.9.1950. Teil I. Heft 3. Textheft. 45p. 2 loose fold. maps. DM12. Teil II. Heft 5. Textheft. 36p. loose fold. map. DM7. Wiesbaden, 1956. 11¾″. (Statistik der Bundesrepublik Deutschland 36, 37)

Die Bevölkerung der Bundesrepublik Deutschland nach der Zählung vom 13.9.1950. Heft 9. Textheft. Wiesbaden, 1956. 117p. 5 loose fold. maps. 11¾″. DM15. (S.B.D. 35)

Gebäude- und Wohnungszählung in der Bundesrepublik Deutschland vom 13. September 1950. Heft 1. Einführung in die Methoden und die Organisation der Zählung. Wiesbaden, 1956. 49p. 11¾″. DM2.50. (S.B.D. 38)

Die Kostenstruktur in der gewerblichen Wirtschaft und in ausgewählten freien Berufen (Ergebnisse der Kostenstrukturerhebung 1950). Heft 1. Industrie. 88p. DM4.50. Heft 3. Grosshandel und Verlagsgewerbe. 114p. DM5.70. Heft 5. Gaststätten- und Beherbergungsgewerbe. 35p. DM1.80. Heft 7. Ärzte und Zahnärzte. 16p. DM.80. Wiesbaden, 1956. 11¾″. (S.B.D. 49)

Die nichtlandwirtschaftlichen Arbeitsstätten in der Bundesrepublik Deutschland nach der Zählung vom 13.9.1950 (Textband mit Kartenwerk). Heft 1. Textband. Wiesbaden, 1956. 120p. 11¾″. DM6. (S.B.D. 47)

Italy. *Istituto Centrale di Statistica.* Elezione del Senato della Repubblica, 7 giugno 1953. Vol. I. Legislazione elettorale—certificati—elettori—votanti—risultati per comune desunti dai verbali elettorali di sezione. Roma, 1956. 443p. 10½″. *L*1500.

Moçambique. *Repartição Tecnica de Estatistica.* Recenseamento geral da população em 1950. III. População não civilizada. Lourenço Marques, 1955. xli, 419, A16p. map. 6 fold. diagrs. 10½″.

New Zealand. *Department of Health. Medical Statistics Branch.* Accounting for sickness and death: a guide for coroners, doctors, registrars, and medical records staff. Wellington, 1955. 48p. 9½″.

Report of the British Empire Cancer Campaign (N.Z. Branch) cancer registration scheme. Wellington, 1954. 118p. 10 tables. 13¾″.

Department of Statistics. Final report on the New Zealand census of farm production, 1949–50. (Conducted as part of the World Census of Agriculture of the Food and Agriculture Organization of the United Nations.) Wellington, 1956. 87p. 12¼″.

Norway. *Statistisk Sentralbyrå.* Folketellingen 1.desember 1950. Tredje hefte . . . Population by principal occupation in the rural and town municipalities and counties. Oslo, 1956. 236p. 9½″. *kr.*4. (Norges Offisielle Statistikk XI, 221)

Organisation for European Economic Co-operation. The coal industry in Europe. Paris, 1956. London, H.M.S.O. 84p. 9½″. 6*s*. (Trends in Economic Sectors)

Demographic trends in Western Europe, 1951–1971: a report by the Manpower Committee. Paris, 1956. 155p. 9½″. 9*s*.

The engineering industries in Europe: a study by the Machinery Committee. Paris, 1956. 207p. 9½″. 10*s*. (Trends in Economic Sectors)

The iron and steel industry in Europe. Paris, 1956. 120p. 9½″. 9*s*. (Trends in Economic Sectors)

Sweden. *Statistiska Centralbyrån.* Jordbruksräkningen 1951. Stockholm, 1956. xxii, 118°, 458p. illus. 9¾″. (Sveriges Officiella Statistik: Jordbruk med Binäringar)

Union of South Africa. *Bureau of Census and Statistics.* Population census, 8th May, 1951. Vol. II. Marital status of the white population of the Union of South Africa, together with 1946 census figures for all races of the population. Pretoria, 1955. ix, 154p. 12¼″. 30*s*. 6*d*.

United Nations. *Department of Economic and Social Affairs. Population Branch.* Analytical bibliography of international migration statistics, selected countries, 1925–1950. New York, 1955. London, H.M.S.O. v, 195p. 11″. 14*s*. (Popul. Stud. 24)

Economic Commission for Europe. Inland Transport Committee. The problem of cost in the inland transport industry: a conspectus of the work of the costing experts; compiled by André Brunet, chairman. [Geneva, 1956.] 2v. 11″.

Economic Commission for Europe. Transport Division. Statistics of road traffic accidents in Europe, 1954. Geneva, 1956. viii, 45p. 11″. 2*s*. First annual report.

Educational, Scientific and Cultural Organization. Books for all: a study of international book trade; by R. E. Barker. Paris, 1956. 102p. 10 col. diagrs. 10¾″. 15*s*.

World Health Organization. Parts of certain reports of the Sub-committee on the registration of cases of cancer as well as their statistical presentation convened by the Executive Board at its fourth session in implementation of a request of the second World Health Assembly. *W.H.O. Tech. Rep. Ser.* (1950), No. 25, (1952), No. 53. 5p. 13½″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

United States. *Children's Bureau.* Survey methods for determining the need for services to children of working mothers; [by] Jack Wiener. Washington, Superintendent of Documents, 1956. [iii], 48p. 9¼″. 25*c.*

II.—AUTHORS AND MISCELLANEOUS

Acton Society Trust. Management succession: the recruitment, selection, training and promotion of managers. London, 1956 [vii], 139p. 9¾″. 10*s.* 6*d.*

Adams, Joe Kennedy. Basic statistical concepts. New York & London, McGraw-Hill, 1955. xvi, 304p. 9¼″. 41*s.* 6*d.*

Adler, Max K. Market and opinion research: their use for newspapers and magazines. *Gazette, Leiden* (1956), **2**. 11p. 9½″.

Armsen, P. Tables for significance tests of 2 × 2 contingency tables. *Biometrika* (1955), **42**, 494–511. 10¾″. 2*s* 6*d.* (New Statist. Tab. 21)

Barna, Tibor. Classification and aggregation in input-output analysis. [Offprint from] *The Structural interdependence of the economy: proceedings of an international conference on input-output analysis, Varenna, 27 June–10 July, 1954.* New York, Wiley. Milan, Giuffrè, 1955. [ii], 175–185p. 9½″.

Introduction. [Offprint from] *The structural interdependence of the economy . . .* New York, Wiley, Milano, Guiffrè, 1955. [ii], 21–38p. 9½″.

Bellerby, J. R. *& others.* Agriculture and industry: relative income . . . London, Macmillan, 1956. xii, 369p. fold. table. 8½″. 30*s.*

Benedetti, Carlo. Di un massimo dell' indice quadratico di oscillazione. *Metron* (1956), **18**, No. 1–2. 20p. 9½″.

Benney, M., Gray, A. P. & Pear, R. H. How people vote: a study of electoral behaviour in Greenwich. London, Routledge & Kegan Paul, 1956. xii, 227p. 8½″. 25*s.*

Booth, Andrew D. & Booth, Kathleen H. V. Automatic digital calculators. 2nd ed. London, Butterworth, 1956. ix, 261p. 3 plates. 8½″. 32*s.*

British Iron and Steel Research Association. *Plant Engineering Division.* Conference on works transport: in the chair, Mr. G. H. T. Williams; held at Ashorne Hill, Leamington Spa, 5th & 6th July, 1955. London, [1956]. 52p. 9¾″. 5*s.* Restricted.

British Market Research Bureau Ltd. Readings in market research: a selection of papers by British authors. London, 1956. xxxvi, 235p. 10″. 35*s.*

British Productivity Council. Freight handling. London, 1956. 47p. photos. 8½″. 2*s.* 6*d.* (B.P.C. Productivity Review 24)

Policy and progress, 1955–56. London, 1956. 36p. 9¼″. 1*s.*

Cambridge. University. *Department of Applied Economics.* The underlying causes of the autumn crisis; by E. A. G. Robinson. *Times Rev. Ind., Lond. & Camb. Econ. Bull.* (1955), No. 16. 8p. 10¼″. (Repr. Ser. 105)

Model building and the social accounts: a survey; by Richard Stone. *Income & Wealth, Ser. IV* (1955), 27–77. 51p. 8½″. (Repr. Ser. 106)

International comparisons of real product: an approach by industry of origin; by Deborah Paige & S. Adler. *Income & Wealth, Ser. IV* (1955), 120–144. 25p. 8½″. (Repr. Ser. 107)

Cornell, Francis G. The essentials of educational statistics. New York, Wiley, 1956. London, Chapman & Hall. xii, 375p. 9″. 46*s.*

Coulter, J. L. S. The Royal Naval Medical Service . . . Vol. II. Operations. London, H.M.S.O., 1956. xvii, 543p. 19 photos. map. 9½″. 57*s.* 6*d.* (History of the Second World War, U.K. Medical Series)

Devons, Ely. An introduction to British economic statistics. Cambridge, University Press, 1956. vii, 255p. 7¼″. 22*s.* 6*d.*

Drion, E. F. Estimation of the parameters of a straight line and of the variances of the variables, if they are both subject to error. Publication of the Statistics Department T.N.O. (Mathematics), [195–], 256–260. 5 l. 10¼″. Communicated by Prof. D. van Dantzig at the meeting of May 26, 1951.

Some distribution-free tests for the difference between two empirical cumulative distribution functions. *Ann. Math. Statist.* (1952), **23**, 563–574. 10″.

Enrick, Norbert Lloyd. Quality control through statistical methods, specifically designed for textile mills. New York, Rayon Publishing Corp., 1954. 72p. 11¼″. $5.

Ente Nazionale Idrocarburi. ENI 1955. Roma, 1956. 103p. photos. illus. 10¼″ × 9½″.

Evans, Ancrum. Cash crisis in agriculture. London, [Massey-Harris-Ferguson], 1956. [24] p. 9″.

Fuel for expansion. Coal; E. F. Schumacher. Oil; Clive Dalton. Electricity and atomic power; Eric Shankleman. *Steel Rev.* (1956), No. 2, 23–40. 20p. 9¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Goldsmith, Raymond W., Brady, Dorothy S. & Mendershausen, Horst. A study of saving in the United States. Vol. III. Special studies. Princeton, University Press, 1956. London, Cumberlege. xix, 476p. 9¾″. 68*s*.

Grant, Eugene L. Statistical quality control. 2nd ed. New York, London, [etc.], McGraw-Hill, 1952. xvi, 557p. 9″. 52*s*. 6*d*.

Graves, Lawrence M. The theory of functions of real variables. New York & London, McGraw-Hill, 1956. xii, 375p. 9″. 56*s*. 6*d*.

Gutierrez, Gustavo. La carta magna de la Comunidad de las Naciones. La Habana, Lex, 1945. [viii], 589p. 8¼″.

Harewood, J. A sytem of labour force statistics. *Social Econ. Stud., Jamaica* (1956), **5**, 1–18. 9½″.

Hauser, Philip M. & Leonard, William R., *editors*. Government statistics for business use. 2nd ed. New York, Wiley, 1956. London, Chapman & Hall. xx, 440p. 9″. 68*s*.

Heyworth, *Sir* Geoffrey. The managers. [London], Unilever, 1956. [ii], 18p. 8¼″. The speech delivered at the Annual Meeting of Unilever, 1956, reproduced for private circulation.

Hickman, C. Addison & Kuhn, Manford H. Individuals, groups, and economic behaviour. New York, Dryden Press, 1956. xvii, 266p. 8½″. $4.75.

Hicks, J. R. A revision of demand theory. Oxford, Clarendon Press, 1956. ix, 196p. 7¼″. 18*s*.

Hill, A. Bradford & Doll, R. Lung cancer and tobacco: the B.M.J.'s questions answered. *Brit. Med. J.* (1956), **1**, p. 1160. 12p. 8½″.

Hughes, *Mrs.* H. & Joseph, M. E. The market for television in Australia. [Melbourne], Australia and New Zealand Bank Ltd., [1956]. 15p. 9½″.

Illinois. University. *Bureau of Economic and Business Research.* Determinants of capital expenditures: an interview study; by Robert Eisner. Urbana, University of Illinois Press, 1956. 100p. 9″. $1.50. (Stud. Busin. Expectations Plann. 2)

International Committee for Stage-grouping in Cancer and for the Presentation of the Results of Treatment of Cancer. First report of the International Committee . . . (I.C.P.R.) to the International Congress of Radiology. Copenhagen, 1953. 14p. 13¾″.

Jerusalem. Hebrew University. *Eliezer Kaplan School of Economics and Social Sciences.* Studies in economics and social sciences; edited . . . by Roberto Bachi. Jerusalem, Magnes Press, 1956. [vi], 396, [2]p. 9¾″. (Publ. Hebr. Univ., Jerusalem, Vol. 3)

Keen, Joan E. Sampling inspection schemes. *Target* (1956), **9**, No. 4, p. 2. 17″ × 12¼″.

Lehoczky, Zygmunt. Childbirth, complications and mortality of obstetrical patients in British Columbia (Health Statistics Research Survey). Victoria (B.C.), 1955. x, [118] l. 14¼″.

Lloyd, E. M. H. Fresh eggs and free markets: an open letter to Sir James Turner, President of the National Farmers' Union and the other promoters of the Egg Marketing Scheme. London, The Society of Objectors to Compulsory Egg Marketing, 1956. 10p. 8¼″. 6*d*.

Unscrambling the egg. *Daily Telegraph* (1956), March 17th. 1p. 10″.

London. Stock Exchange. Interest and dividends upon securities quoted on the Stock Exchange, London, 1955. London, 1956. 23p. 10″.

London. University. *University College.* Studies in communication; contributed to the Communication Research Centre, University College, London . . . London, Secker & Warburg, 1955. vii, 182p. 10 plates. illus. 8½″. 21*s*.

Melman, Seymour. Dynamic factors in industrial productivity. Oxford, Blackwell, 1956. xiii, 238p. 8½″. 22*s*. 6*d*.

Milk Marketing Board. The work of the Milk Marketing Board of England and Wales. [Thames Ditton, 1956.] 46p. photos. 8½″.

Mols, Roger. Introduction à la démographie historique des villes d'Europe du XIVe au XVIIIe siècle. Gembloux (Belgium), Duculot, 1954–56. 3 v. 10″. *fr*.1100. (Université de Louvain: Recueil de Travaux d'Histoire et de Philologie. Sér. 4, Fasc. 1)

Myburgh, C. A. L. A brief comparison of the fertility and mortality rates of Africans in various countries. *Cent. Afr. J. Med.* (1956), **2**, 155–159. 9¾″.

National Bureau of Economic Research. Basic research and the analysis of current business conditions; [by] Solomon Fabricant, Director of Research. 36th annual report: a record of 1955 and plans for 1956. New York, 1956. London, Oxford University Press. [vi], 88p. 10¾″.

Consumption and business fluctuations: a case study of the shoe, leather, hide sequence; [by] Ruth P. Mack. New York, N.B.E.R., 1956. xvii, 293p. 11″. $7.50. (Stud. Busin. Cycles 7)

Resource and output trends in the United States since 1870; [by] Moses Abramovitz. New York, N.B.E.R., 1956. 23p. 9¼″. 50*c*. (Occ. Pap. 52)

National Coal Board. Investing in coal: progress and prospects under the plan for coal. London, N.C.B., 1956. 23p. 11″. 3*s*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Netherlands. *Centraal Planbureau.* Vermongensvorming door de staat; door W. M. N. van den Wildenberg. *Ann. Thijmgenootschap* (1955), **43**, 296–315. 20p. 9½″. (Overdrukken 43)

Enige ramingen in verband met het huurvraagstuk; door Dr. W. Drees, Jr. *Econ.-statist. Ber.* (1956), **41**, 253–256. 11p. 9½″. (Overdrukken 46)

Toekomstige werkgelegenheid voor en aanbod van academisch gevormde chemici; door K. van Dalfsen & F. Th. van der Maden. *Chem. Weekbl.* (1956), **52**, 294–297. 9½″. (Overdrukken 47)

Nordisk statistisk skriftserie 3. De nordiske statistikermøter i Oslo, juni 1954. The statistical conferences of the Northern Countries, Oslo, June 1954. Oslo, Indre Smaalenen, 1955. 185p. 9¼″.

Oslo. Universitet. *Socialøkonomiske Institutt.* The Neyman-Pearson theory of testing hypothesis . . . by R. Frisch. Memor. 24.11.47. 3p. 11¾″.

The cleavings of a univariate measurement into two factors; by R. Frisch & E. Sverdrup. Memor. 23.5.48. 7 l. 11¾″.

Converging dyadic approximations by means of partially unknown elements; by R. Frisch. Memor. 2.10.48. 13 l. 11¾″.

Autonomy of economic relations; by R. Frisch, T. Haavelmo, T. C. Koopmans, J. Tinbergen. Memor. 6.11.48. [1], 60 l. 11¾″.

A system of concepts describing the economic circulation and production process; by O. Aukrust, P. J. Bjerve, R. Frisch. 2nd print. Memor. 12.2.49. 57p. 11¾″.

Repercussion studies at Oslo; by R. Frisch. Memor. 20.4.49. [*Amer. Econ. Rev.* (1948), **38**, 367–372.] 6 l. 11¾″.

The case for planning; by L. R. Klein. 2nd printing. Memor. 30.8.49. 30 l. 11¾″.

Attempt at clarification of certain national income concepts; by R. Frisch. Memor. 8.10.49. 14 l. 11¾″

On the indentifiability of parameters in Thurstone's *Multiple factor analysis*; by O. Reiersøl. Memor. 11.10.49. 25 l. 11¾″.

Axiomatic remarks on some national income concepts; by R. Frisch. Memor. 16.10.49. 11 l. 11¾″.

On the notion of equilibrium and disequilibrium; by R. Frisch. Memor. 4.2.50. [*Rev. Econ. Stud.* (1936), **3**, 100–105.] 7 l. 11¾″.

A summary of some of Professor Haavelmo's publications; by R. Frisch & H. J. Kreyberg. Memor. 11.4.50. 4 l. 11¾″.

Sequential method of sampling for deciding between two courses of action; by A. Wald. Memor. 4.2.51. [*J. Amer. Statist. Ass.* (1945), **40**, 277–306.] 30 l. 11¾″.

The elements of Frisch's *Confluence analysis*; by T. Haavelmo & H. Staehle. Memor. 16.2.51. 36 l. 11¾″.

Monopole—polypole—la notion de force dans l'économie; par R. Frisch. Memor. 12.3.51. [*Nationalökon. Tidskr.* (1933), **71** (Suppl. Til Harald Westergaard), 241–259.] 16 l. 11¾″.

Quantitative research in agricultural economics: the interdependence between agriculture and the national economy; by T. Haavelmo. Memor. 20.3.51. 13 l. 11¾″.

Correlation and scatter in statistical variables; by R. Frisch. Memor. 8.5.51. [*Nord. statist. Tidskr.* (1929), **8**, 36–102.] [2], 62 l. 11¾″.

Investment and consumption as factors in economic cycles; by R. Frisch. Memor. 26.9.51. 6 l. 11¾″.

Knut Wicksell: a cornerstone in modern economic theory; by R. Frisch. Memor. 15.12.51. [2], 46 l. 11¾″.

A bibliography of Knut Wicksell's published works; by A. Amundsen. Memor. 12.3.52. 7 l. 11¾″.

An axiomatization of reasonable behaviour in the face of uncertainty; by L. J. Savage. Memor. 13.3.52. 5 l. 11¾″.

The occurrence test: by R. Frisch . . . Memor. 25.4.52. [1], 52 l. 11¾″.

Ricardo's views on monetary questions; by R. S. Sayers. Memor. 30.4.52. 18 l. 11¾″.

Annual survey of general economic theory: the problem of index numbers; by R. Frisch. Memor. 3.10.52. [*Econometrica* (1936), **4**, 1–38.] 12 l. 11¾″.

How can we decide whether the minimax principle is good or bad? by E. Sverdrup. Memor. 12.12.52. 7 l. 11¾″.

Price—wage—tax—subsidy policies as instruments in maintaining optimal employment: a memorandum on analytical machinery to be used in discussions on causes of and remedies to unemployment; by R. Frisch. Memor. 12.2.53. 65 l. 11¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Some basic principles of price of living measurements; by R. Frisch. Memor. 25.6.53. 19 l. 11¾″.

On welfare theory and Pareto regions: presented as an appendix to a paper "From national accounts to macro-economic decision models" to be read at the meeting of the International Association for Research in Income and Wealth, Rome, September, 1953; by R. Frisch. Memor. 17.8.53. [ii], 65 l. 11¾″. English MS of a paper translated into French for publ. in *Economie appliquée*, Paris.

Sur un problème d'économie pure; par R. Frisch. Memor. 3.5.54. [*Norsk mat. Foren. Skr.* (1926), Ser. 1, No. 16.] [ii], 35 l. 11¾″.

The fundamental assumption of Mr. Keynes' monetary theory of unemployment; by W. W. Leontief. Memor. 13.5.54. 6 l. 11½″.

The rate of interest and the optimum propensity to consume; by O. Lange. Memor. 15.5.54. 21 l. 11¾″.

A scheme for collecting and utilizing data on the Indian economy; by M. Mukherjee, with a foreword by Ragnar Frisch. Memor. 18.6.54. [ii], 18 l. 11¾″.

Analysis of binomial experiments; by O. Reiersøl. Memor. 2.8.54. 25 l. 11¾″.

Linear expenditure functions: an expository paper; by R. Frisch. Memor. 5.8.54. 8 l. 11¾″.

Statistical versus theoretical relations in economic macrodynamics; by R. Frisch. Memor. 10.9.54. 16 l. 11¾″.

A reconsideration of confluence analysis; by O. Reiersøl. Memor. 28.1.53. [i], 54 l. 11¾″.

The logarithmic potential method of convex programming, with particular application to the dynamics of planning for national development: synopsis of a communication to be presented at the International Colloquium of Econometrics in Paris, 23–28 May, 1955; by R. Frisch . . . Memor. 13.5.55. [ii], 35 l. 11¾″.

The multifax method for linear programming; by R. Frisch . . . Memor. 17.10.55. [iv], 49 l. 11¾″.

Pearce, S. C. Experimentation with perennial plants. Rome, F. A. O., [1955]. 2, 51p. 11″. (Training Center in Experimental Designs and Survey Techniques of Experimentation, New Delhi, India, September–December 1954)

Priestley, M. B. & Samet, P. A. The analysis of stationary time series. I. Computation of correlation coefficients on a high speed computer. London, Ministry of Supply, 1956. 10p. 13″. (R. A. E., Farnborough. Tech. Note M.S.27)

Roy, Purnendu Mohon. Analysis of $p \times (p - l)$, n-ple latinized rectangular lattices and their multiples. *Calcutta Statist. Ass. Bull.* (1955), **6**, 113–131. 9¾″.

Staples, Ronald, *editor.* "Taxation" key to income tax and surtax, 1956–57. Budget edition, 1956. London, Taxation Publishing Co., 1956. 223p. 8½″. 10*s.*

Summary of statutory and other requirements in the production of annual accounts of companies; [prepared by a firm of Chartered Accountants]. London, Gee & Co., 1956. 43p. 13″. 7*s.* 6*d.*

Tate, R. F. Applications of correlation models for biserial data. *J. Amer. Statist. Ass.* (1955), **50**, 1078–1095. 9¼″.

Times Literary Supplement. The dynamic society. London, Times Publishing Co., 1956. 16p. 8½″. 6*d.* From the *Times Lit. Supp.* (1956), Feb. 24.

Tokyo. *Institute of Statistical Mathematics.* The Institute of Statistical Mathematics. Tokyo, 1956. 21p. photos. 8¼″.

Vinci, Felice. Il nostro schema di sviluppo economico alla luce dell'esperienza internazionale: con un'appendice epistolare di Vilfredo Pareto. Milano, Unione Tipographica, 1956. 91p. 9¾″. *L.*1000. (Inst. Sci. econ. statist. Opere generali V)

Wellington Harbour Board. Official handbook, containing information relative to the port of Wellington, New Zealand. Wellington, 1956. 96p. photos. 3 fold. maps. 9¼″.

Wijsman, Robert A. A new approach to the mathematical theory of the action of high-energy radiation on single cells. I. Formulation of the theory. II. Consequences of the theory. *Radiation Res.* (1956), **4**, 257–277. 2 pt. 9½″.

Winslet, V. G. Multiple shop companies: organization and management. 3rd ed. London, Gee & Co., 1956. 136p. 8½″. 21*s.*

Wolstenholme, G. E. W. & Millar, Elaine C. P., *editors.* Ciba Foundation symposium on extra-sensory perception . . . London, Churchill, 1956. ix, 240p. 8″. 27*s.* 6*d.*

Yugoslav Statistical Society. III godišnji sastanak . . . Zagreb, 17 do 21 novembra 1955 godine. Third annual meeting. Beograd, [1956]. 2 v. 9½″. Summaries in English or French to papers on Descriptive Statistics. Formerly included in *Statistička Revija.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REGISTRATION OF THE UNITED KINGDOM AND EIRE

No. I.—ENGLAND AND WALES

BIRTHS, DEATHS *and* MARRIAGES *registered in the Calendar years* 1951–1955 *and in the Quarters of those years. Numbers, Annual and Quarterly Rates* per* 1,000 *persons living. (Deaths under* 1 *year of age; rate per* 1,000 *related Live Births. Stillbirths per* 1,000 *births).*

Years	1951		1952		1953		1954		1955†	
Estimated Mid-year Home Popn. in thousands*	43,815(*a*)		43,955(*a*)		44,109(*a*)		44,274		44,441	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
Live births	679,689	*15·5*	673,286	*15·3*	681,964	*15·5*	672,635	*15·2*	664,711	*15·0*
Stillbirths	16,019	*23·0*	15,647	*22·7*	15,683	*22·5*	16,202	*23·5*	15,748	*23·1*
Deaths	549,380	*12·5*	497,484	*11·3*	503,529	*11·4*	501,896	*11·3*	518,657	*11·7*
Marriages	360,624	*8·2*	349,308	*8·0*	344,998	*7·8*	341,731	*7·7*	357,431	*8·2*
Infant Mortality	20,223	*29·7*	18,555	*27·6*	18,324	*26·8*	17,160	*25·4*	16,515	*24·9*
Effective Reproduction Rate	*1·001*		*1·009*		*1·035*		—(*b*)		—(*b*)	
Quarters	Live Births in the Quarters of each Calendar Year									
Jan.–March	176,047	*16·3*	173,338	*15·9*	171,649	*15·8*	173,547	*15·9*	169,361	*15·5*
Apr.–June	181,614	*16·6*	174,159	*15·9*	179,477	*16·3*	175,290	*15·9*	171,191	*15·5*
July–Sept.	168,033	*15·2*	168,030	*15·2*	172,943	*15·6*	166,938	*15·0*	166,584	*14·9*
Oct.–Dec.	153,995	*13·9*	157,759	*14·3*	157,895	*14·2*	156,860	*14·1*	157,575	*14·1*
	Stillbirths									
Jan.–March	4,259	*23·6*	4,174	*23·5*	3,975	*22·6*	4,236	*23·8*	3,997	*23·1*
Apr.–June	4,196	*22·6*	4,004	*22·5*	4,220	*23·0*	4,171	*23·2*	4,017	*22·9*
July–Sept.	3,752	*21·8*	3,714	*21·6*	3,711	*21·0*	3,939	*23·1*	3,897	*22·9*
Oct.–Dec.	3,812	*24·2*	3,755	*23·2*	3,777	*23·4*	3,856	*24·0*	3,837	*23·8*
	Deaths (excluding Stillbirths)									
Jan.–March	205,878	*19·1*	146,948	*13·4*	171,350	*15·8*	153,120	*14·0*	168,628	*15·4*
April–June	121,683	*11·1*	115,369	*10·6*	114,644	*10·4*	117,205	*10·6*	124,230	*11·2*
July–Sept.	100,057	*9·1*	98,613	*8·9*	98,545	*8·9*	104,114	*9·3*	101,680	*9·1*
Oct.–Dec.	121,762	*11·0*	136,554	*12·4*	118,990	*10·7*	127,457	*11·4*	124,119	*11·1*
	Marriages									
Jan.–March	109,567	*10·2*	106,629	*9·8*	93,610	*8·6*	95,922	*8·8*	100,696	*9·2*
April–June	66,382	*6·1*	69,489	*6·4*	77,123	*7·0*	74,364	*6·8*	74,763	*6·8*
July–Sept.	111,312	*10·1*	103,411	*9·4*	105,614	*9·5*	104,505	*9·4*	106,217	*9·5*
Oct.–Dec.	73,363	*6·7*	69,779	*6·3*	68,651	*6·2*	66,940	*6·0*	75,755	*6·8*
	Infant Mortality									
Jan.–March	6,572	*38·4*	5,629	*33·6*	5,709	*34·0*	5,256	*31·0*	4,824	*29·0*
April–June	5,066	*28·6*	4,332	*25·1*	4,448	*25·2*	4,304	*24·7*	4,125	*24·4*
July–Sept.	4,003	*23·4*	3,863	*22·8*	3,862	*22·2*	3,543	*21·0*	3,605	*21·6*
Oct.–Dec.	4,582	*28·5*	4,731	*29·0*	4,305	*26·2*	4,057	*25·1*	3,961	*24·7*

* Home population comprises the resident civilian population plus any British, Commonwealth or Allied Armed Forces stationed in the country. Birth, death and marriage rates are calculated accordingly.
† Provisional figures
(*a*) Revised in the light of final data from the 1951 Census.
(*b*) Not yet available.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

No. II.—SCOTLAND

BIRTHS, DEATHS *and* MARRIAGES *registered in the Calendar years* 1951–1955 *and in the Quarters of those years. Numbers, Annual and Quarterly Rates* per* 1,000 *persons living. (Deaths under* 1 *year of age, rate per* 1,000 *Live Births; Stillbirths per* 1,000 *Births).*

Years	1951		1952		1953		1954		1955†	
Estimated Mid-year Popn. in thousands*	5,114		5,114		5,118		5,123		5,133	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
Live Births	90,639	*17·7*	90,422	*17·7*	90,913	*17·8*	92,315	*18·0*	92,535	*18·0*
Stillbirths	2,479	*27*	2,430	*26*	2,308	*25*	2,401	*25*	2,331	*25*
Deaths	65,778	*12·9*	61,510	*12·0*	58,878	*11·5*	61,380	*12·0*	61,645	*12·0*
Marriages	41,361	*8·1*	41,154	*8·0*	40,902	*8·0*	41,975	*8·2*	43,207	*8·4*
Infant Mortality	3,391	*37*	3,181	*35*	2,800	*31*	2,861	*31*	2,811	*30*
Quarters	Live Births in the Quarters of each Calendar Year									
Jan.–March	23,625	*18·7*	22,892	*18·0*	22,860	*18·1*	23,456	*18·6*	23,111	*18·3*
April–June	24,310	*19·1*	23,531	*18·5*	23,791	*18·6*	24,097	*18·9*	23,942	*18·7*
July–Sept.	21,932	*17·0*	22,161	*17·2*	22,166	*17·2*	22 823	*17·7*	22,734	*17·6*
Oct.–Dec.	20,772	*16·1*	21,838	*17·0*	22,096	*17·1*	21,939	*17·0*	22,748	*17·6*
	Stillbirths									
Jan.–March	657	*27*	677	*29*	602	*26*	612	*25*	597	*25*
April–June	658	*26*	598	*25*	586	*24*	628	*25*	567	*23*
July–Sept.	575	*26*	541	*24*	561	*25*	601	*26*	598	*26*
Oct.–Dec.	589	*28*	614	*27*	559	*25*	560	*25*	569	*24*
	Deaths (excluding Stillbirths)									
Jan.–March	22,428	*17·8*	18,394	*14·5*	17,264	*13·7*	18,003	*14·3*	19,078	*15·1*
April–June	15,442	*12·1*	14,270	*11·2*	14,345	*11·2*	14,638	*11·5*	14,998	*11·7*
July–Sept.	13,104	*10·2*	12,813	*10·0*	12,722	*9·9*	13,327	*10·3*	2,733	*9·9*
Oct.–Dec.	14,804	*11·5*	16,033	*12·5*	14,547	*11·3*	15,412	*11·9*	14,836	*11·5*
	Marriages									
Jan.–March	11,560	*9·2*	12,420	*9·8*	12,085	*9·6*	12,371	*9·8*	12,424	*9·8*
April–June	8,958	*7·0*	8,156	*6·4*	8,212	*6·4*	8,553	*6·7*	8,662	*6·8*
July–Sept	12,048	*9·3*	11,894	*9·3*	12,069	*9·4*	12,490	*9·7*	12,968	*10·0*
Oct.–Dec.	8,795	*6·8*	8,684	*6·8*	8,536	*6·6*	8,561	*6·6*	9,153	*7·1*
	Infant Mortality									
Jan.–March	1,088	*46*	872	*38*	776	*34*	892	*38*	829	*36*
April–June	850	*35*	801	*34*	696	*29*	728	*30*	692	*29*
July–Sept.	698	*32*	663	*30*	614	*28*	555	*24*	588	*26*
Oct.–Dec.	755	*36*	845	*39*	714	*32*	686	*31*	702	*31*

* The population is taken to be the civil population (including members of the Mercantile Marine for the time being in Scotland) and members of the armed forces stationed in Scotland. Birth, death and marriage rates are calculated accordingly.

† Provisional figures.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

No. III.—NORTHERN IRELAND

BIRTHS, DEATHS *and* MARRIAGES *registered in the Calendar years* 1951–1955 *and in the Quarters of those years. Numbers, Annual and Quarterly Rates* per* 1,000 *persons living. (Deaths under* 1 *year of age, rate per* 1,000 *Live Births).*

Years	1951		1952		1953		1954		1955†	
Estimated Mid-year Home Popn. in thousands*	1,373		1,375		1,384		1,387		1,394	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
Live births	28,477	*20·7*	28,760	*20·9*	28,984	*20·9*	28,803	*20·8*	28,965	*20·8*
Stillbirths[1]			Not registered							
Deaths	17,628	*12·8*	14,812	*10·8*	14,813	*10·7*	15,124	*10·9*	15,407	*11·1*
Marriages	9,414	*6·9*	9,300	*6·8*	9,416	*6·8*	9,154	*6·6*	9,510	*6·8*
Infant Mortality	1,173	*41*	1,117	*39*	1,090	*38*	951	*33*	937	*32*
Quarters	Live Births in the Quarters of each Calendar Year									
Jan.–March	7,254	*21·2*	7,095	*20·8*	7,212	*21·3*	7,253	*21·3*	6,876	*20·2*
April–June	7,769	*22·6*	7,589	*22·2*	7,902	*22·9*	7,704	*22·3*	7,679	*22·1*
July–Sept.	7,089	*20·5*	7,215	*20·9*	7,254	*20·8*	7,089	*20·3*	7,638	*21·7*
Oct.–Dec.	6,365	*18·4*	6,861	*19·9*	6,616	*19·0*	6,757	*19·3*	6,772	*19·3*
	Deaths (excluding Stillbirths)									
Jan.–March	6,971	*20·4*	4,450	*13·0*	4,542	*13·4*	4,311	*12·6*	5,083	*14·9*
April–June	3,951	*11·5*	3,671	*10·7*	3,733	*10·8*	3,985	*11·5*	3,867	*11·1*
July–Sept.	3,211	*9·3*	3,093	*9·0*	3,075	*8·8*	3,197	*9·1*	3,126	*8·9*
Oct.–Dec.	3,495	*10·1*	3,598	*10·4*	3,463	*9·9*	3,631	*10·4*	3,331	*9·5*
	Marriages									
Jan.–March	2,210	*6·5*	2,272	*6·7*	2,022	*6·0*	2,067	*6·1*	2,185	*6·4*
April–June	1,958	*5·7*	2,099	*6·1*	2,288	*6·6*	2,123	*6·1*	2,180	*6·3*
July–Sept.	3,174	*9·0*	2,988	*8·7*	3,235	*9·2*	3,183	*8·9*	3,297	*9·5*
Oct.–Dec.	2,072	*5·9*	1,941	*5·6*	1,871	*5·4*	1,781	*5·0*	1,848	*5·3*
	Infant Mortality									
Jan.–March	371	*51*	317	*45*	332	*46*	285	*39*	268	*39*
April–June	292	*38*	288	*38*	294	*37*	245	*32*	231	*30*
July–Sept.	255	*36*	232	*32*	225	*31*	202	*28*	216	*28*
Oct.–Dec.	255	*40*	280	*41*	239	*36*	219	*32*	222	*33*

* Home population comprises civilians and members of the armed forces stationed in Northern Ireland. Birth, Death and marriage rates are calculated accordingly.
† Provisional figures.
[1] Stillbirths are not registered in Northern Ireland. The birth of one living child and one stillborn, or one living child and two stillborn, is counted as one birth.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

No. IV.—EIRE

BIRTHS, DEATHS *and* **MARRIAGES** *registered in the Calendar years* 1951–1955 *and in the Quarters of those years. Numbers, Annual and Quarterly Rates* per* 1,000 *persons living. (Deaths under* 1 *year of age, rate per* 1,000 *Live Births).*

Years	1951		1952		1953†		1954†		1955†	
Estimated Mid-year Popn. in thousands	2,959		2,948		2,942		2,933		2,909	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
Live births	62,878	21·2	64,631	21·9	62,202	21·1	61,897	21·1	61,575	21·2
Stillbirths‡			Not registered							
Deaths	42,382	14·3	35,105	11·9	34,585	11·8	35,570	12·1	36,525	12·6
Marriages	16,017	5·4	15,876	5·4	15,603	5·3	15,757	5·4	16,137	5·5
Infant Mortality	2,876	46	2,674	41	2,448	39	2,345	38	2,249	37
Quarters	Live Births in the Quarters of each Calendar Year									
Jan.–March	15,727	21·2	16,341	22·2	15,692	21·3	15,278	20·8	15,073	20·6
April–June	16,966	22·9	17,502	23·7	16,761	22·8	16,142	22·0	16,639	22·9
July–Sept.	16,299	22·0	16,048	21·8	15,702	21·3	15,837	21·6	15,674	21·6
Oct.–Dec.	13,886	18·8	14,740	20·0	14,047	19·1	14,640	20·0	14,189	19·5
	Deaths (excluding Stillbirths)									
Jan.–March	16,659	22·5	10,657	14·5	10,555	14·3	10,468	14·2	11,698	16·0
April–June	10,205	13·8	8,713	11·8	8,926	12·1	8,861	12·1	9,699	13·3
July–Sept.	7,584	10·2	7,256	9·8	7,129	9·7	7,708	10·5	7,403	10·2
Oct.–Dec.	7,934	10·7	8,479	11·5	7,975	10·8	8,533	11·6	7,725	10·6
	Marriages									
Jan.–March	3,389	4·6	3,373	4·6	2,935	4·0	3,237	4·4	3,086	4·2
April–June	3,857	5·2	3,754	5·1	3,771	5·1	3,687	5·0	3,866	5·3
July–Sept.	5,437	7·3	5,196	7·1	5,435	7·4	5,530	7·5	5,838	8·0
Oct.–Dec.	3,334	4·5	3,553	4·8	3,462	4·7	3,303	4·5	3,347	4·6
	Infant Mortality									
Jan.–March	928	59	886	54	698	44	743	49	673	45
April–June	783	46	617	35	682	41	589	36	595	36
July–Sept.	596	37	574	36	514	33	469	30	467	30
Oct.–Dec.	569	41	597	41	554	39	544	37	514	36

* Rates are based on the total estimated population as at the middle of the corresponding year.
† Provisional figures.
‡ Stillbirths are not registered in Eire. The births of one living and one stillborn, or of one living child and two stillborn are counted as one birth.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

VOL. 119.] [PART II

Journal of the Royal Statistical Society

SERIES A (GENERAL)

PART II, 1956.

Some Aspects of the Geometry of Statistics

The Use of Visual Presentation in Understanding the Theory and Application of Mathematical Statistics

[The Inaugural Address of the President, Professor E. S. Pearson, C.B.E., delivered to the Royal Statistical Society, on January 18th, 1956]

1. Introduction

The main title which I have chosen for this address was suggested by an incident in the early stages of development of mathematical statistics in England. In the year 1890, Karl Pearson, then Professor of Applied Mathematics at University College, London, was appointed to the part-time professorship in Geometry at Gresham College in the City of London. Whatever may have been the original intentions of Sir Thomas Gresham, the duties of one of his seven professors 300 years after the founding of his College seem to have consisted in giving some dozen end-of-the-day lectures each year to a semi-popular audience. The first 8 lectures which Pearson gave on the "Scope and Methods of Science" were later developed into his book *The Grammar of Science*; these lectures were, however, followed by 12 on "The Geometry of Statistics", 12 on "The Laws of Chance" and finally in 1893–94 by 10 on "The Geometry of Chance".*

It appears from the summaries that the lectures on the Geometry of Statistics were mainly concerned with methods of visual presentation of what may be termed descriptive statistics. Two years later, however, when the elements of a theory of mathematical statistics were already taking shape in Gower Street, with the development of the properties of the binomial and the hypergeometric series, of the normal distribution and of compound and skew frequency curves, the evening lectures given at Gresham College had changed their emphasis. The audiences were encouraged to carry out a variety of sampling experiments whose results were compared with theory, the whole being illustrated by models and wall diagrams, some of which survive with us at University College today.

The graphical approach was a characteristic feature of Pearson's teaching of Applied Mathematics; it appeared in his courses on geometrical drawing and projection, on graphical mechanics and in the drawing office classes which he conducted for engineering students right up till 1911. It was natural that he should extend this approach to Statistics. In the summary to his first lecture on the Geometry of Statistics he noted:

> "Contest of geometry and arithmetic as scientific tools in dealing with physical and social phenomena. Erroneous opinion that geometry is only a means of popular representation; it is a fundamental method of investigating and analysing statistical material".

To-day our statistical theory and the techniques of analysis which we apply to numerical data are based on mathematical models of far greater complexity and refinement than could be imagined 65 years ago. But if the model is more elaborate, it becomes all the more difficult to judge whether it fits the data as a whole. In last February's discussion before the Society on "The Teaching of Mathematical Statistics at University Level", Professor D. G. Champernowne remarked:

* For the syllabuses of these lectures, see E. S. Pearson (1938), Appendix II.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

> ". . . that theory (of mathematical statistics) is solely concerned with working out the properties of the theoretical models—whereas what matters—and what in one sense is most difficult—is deciding what theoretical model best corresponds to the real world situation to which statistical methods must be applied. There is great danger that mathematical pupils will imagine that a knowledge of mathematical statistics alone makes a statistician."

I have little doubt that, where this is possible, a visual survey of the "pattern" of his data provides the statistician who is trained to extract a meaning from diagrams with the quickest method of checking whether the model he proposes to use is likely to be appropriate or not. That understanding can be achieved through visual aids may be regarded as a proposition so obvious that it needs no restatement by me; yet there is much evidence that the average mathematically-trained student who forms to-day the raw material of our classes in mathematical statistics is not too well endowed with visual imagination. How far, one may ask, is this a question of teaching both at school and at university level? Later in the discussion just mentioned Professor M. G. Kendall referred to the question of *how* to teach statistics.

> "University teachers", he said, "are unique among professional men in having no training whatsoever in their profession. They have to learn by experience and some of them never do learn. One would expect that in such circumstances a good deal of thought would have been given to the technique by which statistical ideas should be introduced to students. Perhaps it has, but if so the results have not been made available for general use."

My remarks to-day may be regarded in part as an answer to Professor Kendall's challenge. In broad terms we may accept the thesis that appropriate methods of visual presentation can play an important part in helping the statistician in ways such as these: in understanding the meaning of his arithmetical results; in avoiding mistakes through lack of fit of his models; in saving time; and in providing what is often the best means of making clear his methods of analysis to the non-statistician. But the prestige of arithmetical procedures based on algebraic formulae is deeply entrenched in our lecture courses and our text-books, so that few mathematical statisticians will use to the full their visual faculties unless they are trained to do so. In a matter of this kind one can speak confidently only from personal experience, but I believe that the intellectual stimulus which can come from use of the visual imagination may be very great. In the course of my own statistical training I remember two points at which diagrams formed the key to my understanding of a great tract of territory.

At the very beginning of his lectures on correlation my father would describe the difference between what he termed the physicist's concept of a law of relationship between two variables and the statistician's concept of correlation. He would draw on one part of the blackboard a straight line or curve with a few points lying closely on either side of it. $y = f(x)$ represented the law and the points fell off the curve only through what would be termed errors of observation. Next to this he drew a great belt of dots or crosses suggesting what we now would call a bivariate probability density distribution, and he told us how the statistician's problem was to describe this belt as he found it. One could see in this diagram the meaning of two regression lines and of array distributions which might be described in terms of their standard deviations and their shape. Here, in a picture, was a model of a correlation distribution in which to my mind one could *see* the properties as a whole far more surely than one could extract them from an algebraically expressed function $p(x, y)$.

My second flash of understanding developed from the small diagram on p. 509 of R. A. Fisher's 1915 paper in Vol. 10 of *Biometrika*. While the main part of this paper was concerned with the distribution of the correlation coefficient, this diagram introduced me to that fruitful idea, adequately grasped in terms of 3-dimensional geometry of the relation of a mean, a standard deviation, and Student's ratio to a plane, a cylinder and a cone. One *saw*, also why, when the space was filled with the probability density distribution of "normal" variation with its spherical density contours, then and then only problems of distribution theory became beautifully simple. One could see, too, in geometrical terms, how the methods of the analysis of variance consisted in a skilful rotation of axes, so as to separate and compare the appropriate projections of a vector.

I am told that I have been fortunately endowed with an extra amount of visual imagination, and that it is unfair for me to expect others of my staff or students to be helped by statistical

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

geometry. But while it may be true that most mathematicians cannot see things in pictures, indeed may have a better method of grasping the fit of their models to data than I have, they find difficulty in telling me how they get this overall grasp of the appropriateness of a fit. At any rate I feel justified in setting down some of my own pictures. That even in the field of Pure Mathematics, where no question arises of fitting a model to observational data, the value of visual presentation is recognized, is well illustrated in Hilbert and Cohn-Vosser's *Geometry and the Imagination** (1932). This book is filled with well drawn diagrams illustrating the properties of 3-dimensional curves and surfaces. In the first paragraph to his Preface, Hilbert wrote as follows:

> "In mathematics, as in any scientific research, we find two tendencies present. On the one hand, the tendency towards *abstraction* seeks to crystallize the *logical* relations inherent in the maze of material in a systematic and orderly manner. On the other hand, the tendency towards *intuitive understanding* fosters a more immediate grasp of the objects one studies, a live *rapport* with them, so to speak, which stresses the concrete meaning of their relations."

And later:

> "With the aid of visual imagination we can illuminate the manifold facts and problems of geometry, and beyond this it is possible in many cases to depict the geometric outline of the methods of investigation and proof, without necessarily entering into details connected with the strict definitions of concepts and with the actual calculations."

These passages bring out clearly some of the aspects of the geometry of statistics which I have in mind. The function of visual presentation in fostering in the mathematician a lively understanding of his data; the power of a geometrical outline to help the non-mathematical practitioner in grasping the meaning of the techniques of analysis. In the following sections I shall endeavour to illustrate my theme with a few selected examples. For much of what I shall say I could use as text the introductory paragraph of chapter II of R. A. Fisher's (1925) *Statistical Methods for Research Workers*, which runs as follows:

> "The preliminary examination of most data is facilitated by the use of diagrams. Diagrams prove nothing, but bring outstanding features readily to the eye; they are therefore no substitute for such critical tests as may be applied to the data, but are valuable in suggesting such tests, and in explaining the conclusions founded upon them."

2. Geometry and Teaching

2.1. *Training in the Visual Approach*

I shall later be concerned with stressing the value of diagrams to the practising statistician as companions to arithmetic in extracting the meaning from data, but there is no doubt that their function in teaching is also an important one. This is partly from the help they can give towards understanding mathematical models, by describing their properties in geometric as well as algebraic terms. But if the student is to turn into a statistician who makes full use of diagrams in the analysis of observations he will generally need some training in the visual approach to data. Even if the mathematically-minded student believes that he can himself dispense with any form of visual aid, he may regret later in his career that he has not given more thought to what may prove the surest means of communication open to him with some of his non-mathematical clients or colleagues.

Many openings can be found for encouraging the use of the visual imagination in connection with the common distributions and tests used in statistical sampling theory. I shall start by taking two illustrations which concern Student's t-distribution.

2.2. *The Difference Between Two Means*

Authorities have differed widely in the philosophy of their approach to tests of significance. Fig. 1 presents a series of steps in the development of one possible train of reasoning by which we

* A translation was issued by the Chelsea Publishing Company, New York, in 1952 from which I have quoted.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

apply the *t*-test to draw conclusions about the significance of the difference between two mean values.

At the top we see two population distributions to each of which a normal curve has been fitted. The variable is the depth of sapwood in inches measured in telephone poles supplied by two companies*; below these two distributions the dots represent two random samples of 10 observations, based on measurements of poles from each source of supply. We ask: "would these 20

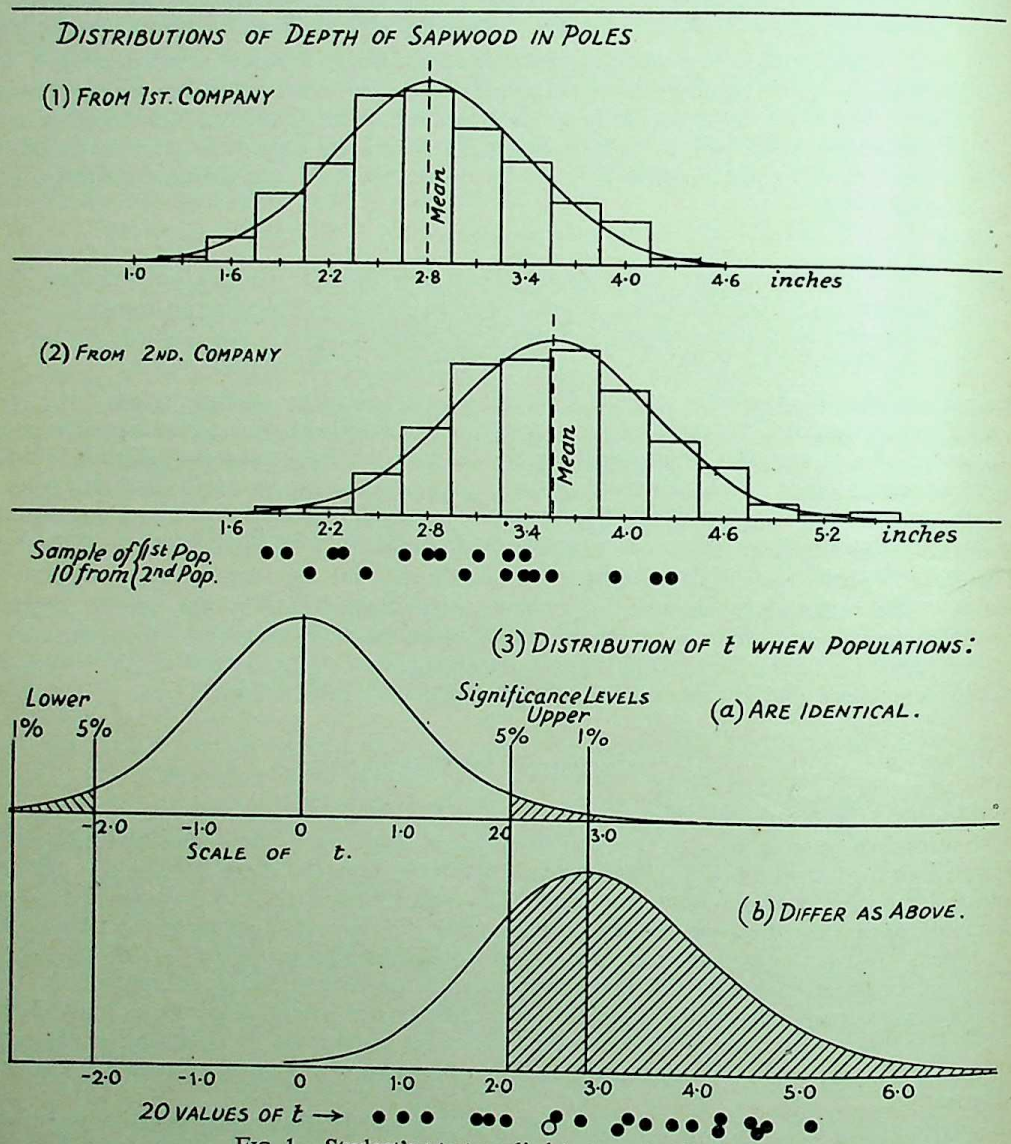


FIG. 1.—Student's *t*-test applied to two sample means.

observations alone give us reason for doubting that the means in the two populations are the same?"

Assuming that the variability in the two populations is approximately equal—and the full data shows this to be the case—we calculate Student's *t*, that is the ratio of the difference between the sample means to the pooled estimate of the standard error of the difference. For this particular pair of samples, $t = 2 \cdot 53$.

* The figures are Bell Telephone Company data originally used by Dr. W. A. Shewhart.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In the next section of the diagram this value of 2·53 is related to the distribution which the t-ratio would follow if the population means were the same; it is seen that it lies between the 5 per cent. and 1 per cent. points (using the two-tailed test with 18 degrees of freedom). In this case two samples of 10 observations would have been sufficient to make the statistician suspect a difference in population means, but it is natural to ask: "would this generally be so?". In other words, given two populations with the degree of overlap illustrated in the top section of the diagram [(Difference between population means)/(standard deviation) = 0·76/0·60 = 1·27] what is the probability that the t-test based on two samples of 10 will establish a difference in means, using either (*a*) the 5 per cent., or (*b*) the 1 per cent. level of significance?

The answer is given mathematically by the so-called non-central distribution of t shown as the lowest curve. Tables or charts (e.g. Pearson and Hartley, 1954, Table 10) indicate that the area under this curve lying outside the 5 per cent. and 1 per cent. limits of the standard t-curve drawn above it are 0·76 and 0·49 respectively. In other words, the probability of establishing significance is (*a*) 0·76, (*b*) 0·49.

To drive home the meaning of this argument pictorially, 19 further pairs of samples of 10 were drawn randomly from the two populations and the 20 values of Student's t (the first value of 2·53 is drawn as an open circle) have been plotted at the foot of the diagram; 14 out of 20 fall beyond the 5 per cent. significance level, and 11 out of 20 beyond the 1 per cent., results in good agreement with the expectations of mathematical theory, i.e. 15·2 and 9·8.

This argument is, of course, very familiar to statisticians, but I have purposely set it out at some length in order to show how Fig. 1, which could easily be reproduced as a class wall diagram, summarizes the underlying properties of the test. Not only so, but it provides points of departure for comment on the adequacy of the mathematical model and the ways in which it may fail. What will happen, it may be asked, if the two population distributions are not well represented by normal curves or have very different variances? How much will the central and non-central t-distributions shown below be modified? Upon this depends what Box has termed the robustness of the test.

There can be little doubt that summary diagrams of this kind have their value for the mathematical statistician in training as well as for the non-mathematician, who besides using tables of significance levels, is seeking for a broad understanding of statistical ideas.

2.3. *The Geometry of Two and Three Dimensions*

My second figure relates to a situation where the visual imagination, extended through two to three dimensions, can help the geometrically-minded to "understand" in a flash what can only be proved algebraically at considerable length. It follows from R. A. Fisher's geometrical identification of the mean and standard deviation (1915, p. 509) already referred to and relates again to a form of Student's t-test.

Suppose that x_1, x_2—and for 3-dimensional representation x_1, x_2, x_3—represents a random sample of 2 (or 3) observations from a normal population, with mean zero and standard deviation σ. Possible samples will be represented by a cluster of points surrounding the origin, 0, in 2 (or 3) dimensions such as are shown in Fig. 2a. Since the probability density law takes the form

$$p(x_1, x_2) = \frac{1}{2\pi\sigma^2}\, e^{-\frac{1}{2}(x_1^2+x_2^2)/\sigma^2}$$

it follows that the contours of this probability distribution are circular or, in 3 dimensions, spherical. If the value of σ is unknown, the probability contours will remain the same, but we shall not know what may be termed the scale of radial concentration about 0.

Suppose that we are uncertain whether the population mean of x is zero, thinking that it may have some unknown positive value, $\mu > 0$. If this is so, in two dimensions, the sample point (x_1, x_2) may be described as one of a possible cluster centred at (μ, μ) on the axis $x_1 = x_2$, for which the density contours will also be circular. This situation again is suggested in Fig. 2a. In the approach commonly followed for testing the hypothesis that $\mu = 0$, on the basis of a given sample, having the alternative $\mu > 0$ in mind, we specify a rejection region in the space such that:

(a) if the observed sample point falls within this region, we reject the hypothesis;
(b) the probability of the point falling in this region, if $\mu = 0$, has some small value α.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

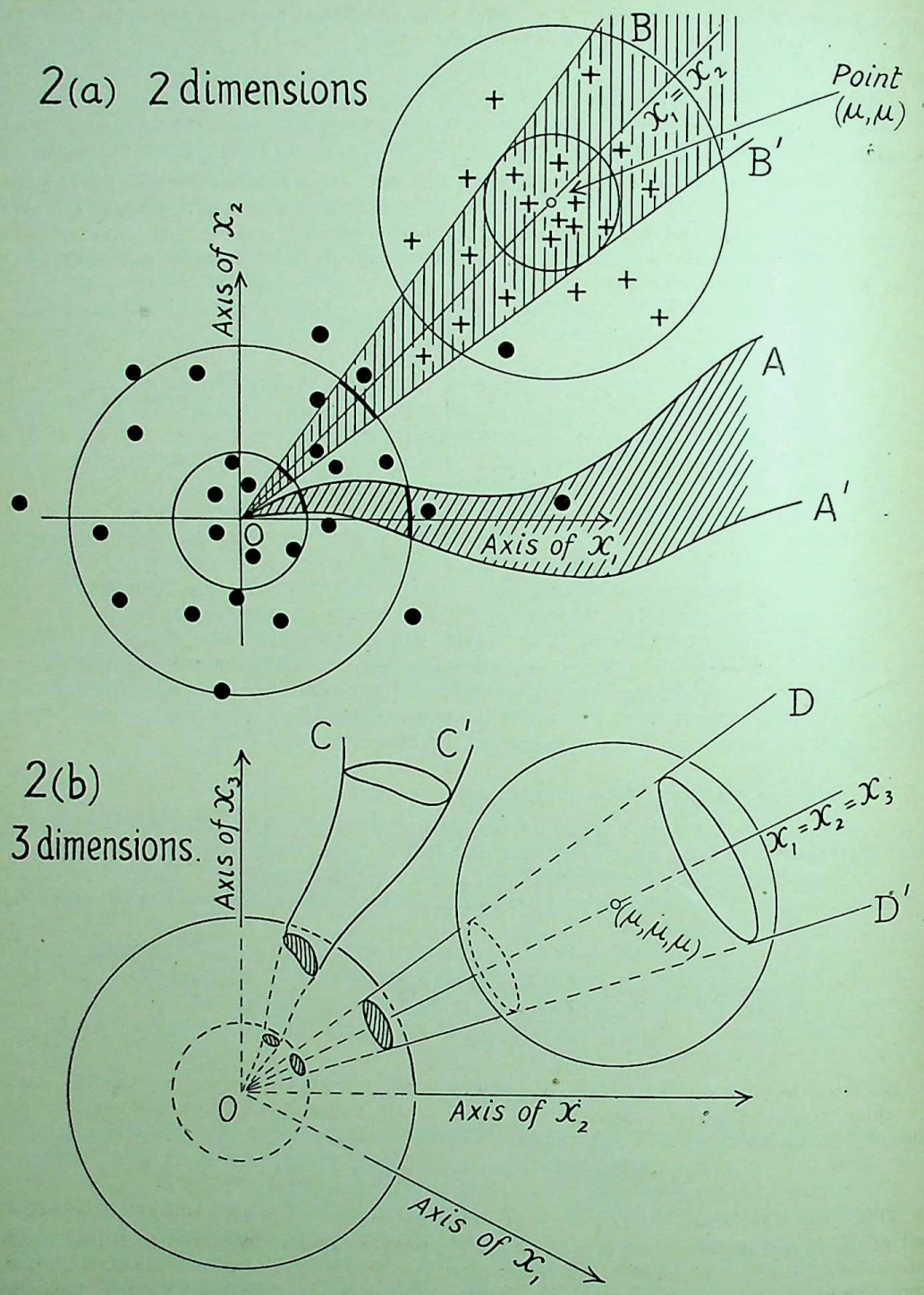


FIG. 2.—Geometric presentation of Student's test.

Now if σ is unknown, geometrical considerations show that the rejection region can only satisfy (b), whatever be σ, if it is built up of elementary regions each containing a fraction α of every circular (or for 3 dimensions, spherical) contour of equal density centred at 0. The horn-shaped

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

regions suggested by the lines AOA′ and COC′ for 2 and 3 dimensions in Fig. 2 are constructed in this way. But how shall we choose between the infinite number of regions which would satisfy requirement (b)?

It appears desirable to choose that region which

(c) will give the greatest chance of rejecting the hypothesis $\mu = 0$ when in fact μ has any other value greater than zero.

Since for $\mu > 0$, the probability density distribution has circular (or spherical) symmetry about some point on the diagonal axis $x_1 = x_2$ (or $x_1 = x_2 = x_3$), it appears to be intuitively established that the rejection region satisfying condition (c) must be directed symmetrically up this diagonal axis—as the area between two straight lines BOB′ in Fig. 2a, or as the volume inside the cone DOD′ with circular cross section in Fig. 2b.

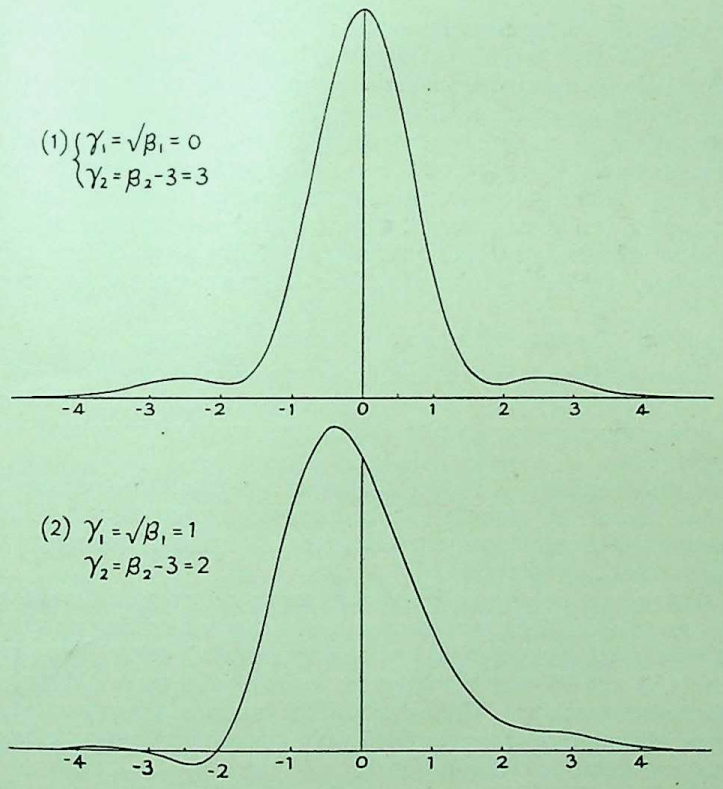


FIG. 3.—Edgeworth curves.

But from Fisher's geometrical result the boundaries of these regions are precisely the surfaces on which Student's t-ratio is constant and *within* which t exceeds the boundary value. For the case of 3 dimensions

$$t = \sqrt{6}\,\bar{x}/\sqrt{\{\Sigma(x_i - \bar{x})^2\}} \qquad \text{where } \bar{x} = \tfrac{1}{3}(x_1 + x_2 + x_3).$$

This result, that Student's ratio provides what has been termed the uniformly most powerful test of the hypothesis specified, can of course be proved in general by algebraic methods. But I am doubtful whether the more conventional proof can provide the same mental satisfaction as an approach by which we seem to see the result must be true by an appeal to our visual imagination!

2.4. *The Shape of Edgeworth Curves*

The previous examples have been concerned with the interpretation of statistical tests in terms of the concept of repeated sampling. My next example is of a different character.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In recent years it has been common to use Gram-Charlier and Edgeworth curves to represent the distribution of the variable in the parent population when exploring the effect of departure from normality on standard statistical procedures, such as those of the analysis of variance. This has been done because the use of the mathematical forms involved renders algebraic treatment possible, even if sometimes laborious. Thus it is assumed that an Edgeworth curve of the form

$$f(x) = e^{-\frac{1}{2}x^2}/\sqrt{(2\pi)}\left\{1 + \frac{1}{3!}\gamma_1 H_3(x) + \frac{1}{4!}\gamma_2 H_4(x) + \frac{1}{3\,.\,4!}\gamma_1^2 H_6(x)\right\}$$

can be used to represent the population, where

$$\gamma_1 = \sqrt{\beta_1} = \kappa_3/\kappa_2^{\frac{3}{2}}, \qquad \gamma_2 = \beta_2 - 3 = \kappa_4/\kappa_2^2$$

are the usual moment ratios of the distribution and $H_3(x)$, $H_4(x)$ and $H_6(x)$ are Hermite polynomials in x.

But there has not always been evidence in the published papers that the authors have examined the shapes of the distributions falling within the range of the γ_1, γ_2 values which they have discussed. Thus the two curves of Fig. 3 have γ-values (0, 3 and 1, 2) which have been included within the range of some of these investigations. It is probable that certain broad conclusions about the robustness of tests can be drawn by stretching departure from normality into shapes containing small negative frequencies and subsidiary bumps, but if the writers had been trained always to *look* at the shapes of frequency distributions and probability curves before starting to use them, they would hardly have included such forms in their discussion without some comment.

2.5. *Diagrams showing Component Parts in the Analysis of Variance*

One of the statistical fields in which visual aids to understanding have always seemed to me valuable is the analysis of variance. Here, the arithmetical concept of the partition of a total sum of squares into component parts which can be measured and compared, may be usefully reinforced by geometry.

My example is taken from O. L. Davies' *Statistical Methods in Research and Production* (1947, pp. 90–93). I shall discuss this example and the following one at some length to illustrate what I mean by that interplay of algebra, geometry and arithmetic which I believe can be used as a means of increasing the student's comprehension.

The data relate to a balanced experiment giving tests on compressive strength of cement mortar, cast into ½-inch cube moulds. The problem was to investigate how far a personal element entered into the results, due to the particular men who either mixed or tested the material. Three men whom I shall call I, II and III were employed as "mixers" and three different men, A, B and C were used on the testing machine. The balance in the experiment was secured by arranging that each of the nine possible mixer-tester pairs dealt with four cubes of cement mortar. The resulting 36 measurements are plotted as circles in Fig. 4.

The mathematical model which may here be introduced into the analysis is as follows. y_{mti} represents the breaking strength of the ith cube ($i = 1, \ldots, 4$) dealt with by mixer m and tester t($m, t = 1, 2, 3$). Then an additive relation is assumed such that

$$y_{mti} = \alpha + \beta_m + \gamma_t + \delta_{mt} + z_{mti} \qquad (1)$$

where α represents a general average level of strength, β_m and γ_t are terms due to individual mixer and tester bias, and δ_{mt} are the so-called interaction terms which, if present, measure the amount by which the combination of mixer m and tester t cannot be represented by the sum of two independent terms, β_m and γ_t . z_{mti} is a residual "error" term which cannot be accounted for by systematic biases.

As a counterpart to the idealized equation (1) we can form an identity with estimated components*

* Here the mean values have their customary meaning, i.e. $y_{\ldots}$ is the grand mean of the 36 observations, $y_{m..}$ the mean of the 12 cubes prepared by mixer m, $y_{.t.}$ of those broken by tester t, and $y_{mt.}$ of the 4 cubes dealt with by m and t in combination.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$y_{mti} \equiv \underset{(\alpha)}{y_{...}} + \underset{(\beta_m)}{(y_{m..} - y_{...})} + \underset{(\gamma_t)}{(y_{.t.} - y_{...})} + \underset{(\delta_{mt})}{(y_{mt.} - y_{m..} - y_{.t.} + y_{...})} + \underset{(z_{mti})}{(y_{mti} - y_{mt.})} \quad . \quad . \quad . \quad (2)$$

The standard analysis of variance table given in Table 1 provides in arithmetical form the partition of the total sum of squares about the grand mean, represented by the identity

$$\begin{aligned} \Sigma(y_{mti} - y_{...})^2 &\equiv \Sigma(y_{m..} - y_{...})^2 + \Sigma(y_{.t.} - y_{...})^2 \\ S \quad &\equiv \quad S_1 \quad + \quad S_2 \\ &\quad + \Sigma(y_{mt.} - y_{m..} - y_{.t.} + y_{...})^2 + \Sigma(y_{mti} - y_{mt.})^2 \\ &\quad + \quad S_3 \quad + \quad S_4 \end{aligned} \quad . \quad (3)$$

or

TABLE 1

Analysis of Variance for Tests on Portland Cement

Source of Variation	*Sum of Squares*	*Degrees of Freedom*	*Mean Square*	*Variance Ratio*
Between mixers	$S_1 = 8{,}965$	2	4,482	$F_1 = \frac{S_1}{2} \div \frac{S_4}{27} = 1{\cdot}63$
Between testers	$S_2 = 25{,}061$	2	12,530	$F_2 = \frac{S_2}{2} \div \frac{S_4}{27} = 4{\cdot}56$
Mixer-tester interaction . . .	$S_3 = 6{,}638$	4	1,659	$F = \frac{S_3}{4} \div \frac{S_4}{27} = 0{\cdot}60$
Residual (within groups of 4 observations)	$S_4 = 74{,}155$	27	2,746	
Total .	$S = 114{,}819$	35		

Inferences are drawn in the usual way by a comparison of the mean squares calculated from this analysis. Thus we find * no evidence for interaction nor for any effect due to the different mixers but there is a definite suggestion of tester bias.

How now can geometry help us in understanding? Fig. 4 illustrates just how the sums of squares set out in equation (3) and Table 1 are made up. Starting from the 36 observational spots we see the mean value for each group of four observations marked by a short horizontal stroke. The residual S_4 of the table is the sum of squares of the 36 distances between each spot and the corresponding column mean. We may now start from the other end and build up towards these nine column means as in equation (2).

First we have the level of the grand mean at $y... = 5012$. On to this we add three steps representing the estimated mixer biases, -7, -15 and 22, drawn as the solid-line components. On to these are added the three estimated tester biases, 24, 13 and -37, which are drawn (three times over) as the hatched-line components.† Do these two sets of additive components, a mixer and a tester term, adequately represent the situation? This will depend on how nearly the ends of the hatched blocks fit on to the nine column means, where "nearness" has to be judged in terms of the residual variation. The nine differences (*viz.* 6, -19, 13; -16, 24, -8; 10, -5, -5) squared, summed and multiplied by four‡ give the interaction sum of squares, S_3 of Table 1.

The visual picture of Fig. 4 cannot by itself provide an answer to the investigation. From it we can see the meaning of the arithmetical analysis—can note, for example, how the three tester biases, estimated as 24, 13 and -37, are more important than the three mixer biases of -7, -15 and 22. But unless our visual imagination could be very highly trained so as to link the impression of the picture with probability tables, we could not safely attempt an assessment of significance. The value of this diagram, however, in giving a clearer meaning to arithmetic and algebra, can surely not be questioned? Whatever conclusions we draw from Table 1 must make

* By referring the variance ratios F shown in Table 1 to appropriate significance tables.
† We could equally well have added the tester biases first and then put the mixer biases on top of these.
‡ Since each of the four observations in a column provides an identical interaction contribution.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sense in the picture of Fig. 4. If they make nonsense, we shall suspect an error in calculation or question whether the model used is appropriate.

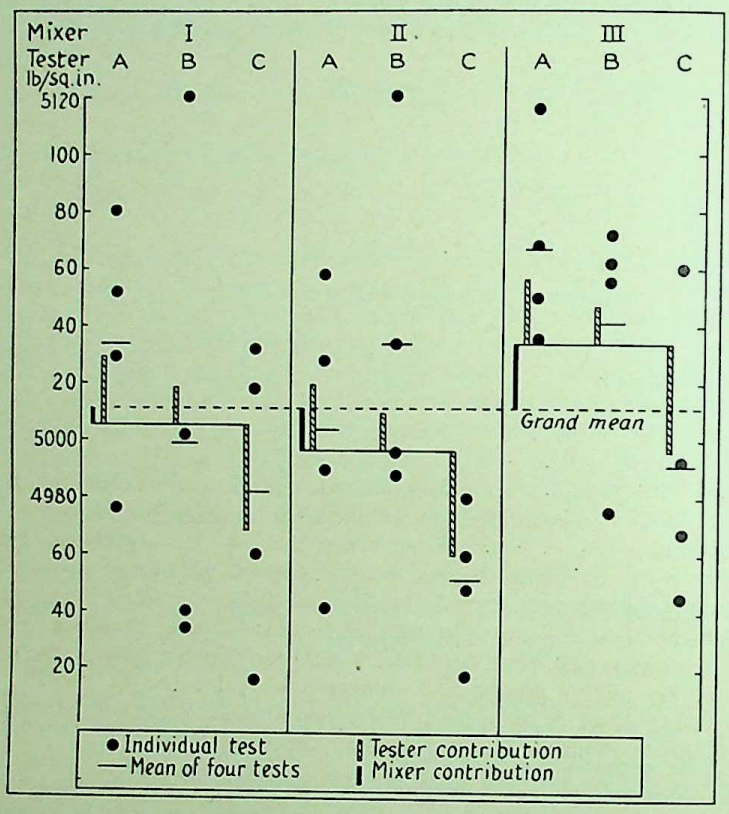


FIG. 4.—Tests on samples of Portland Cement.

2.6. *Diagrams for the Analysis of Covariance*

An allied type of problem in which visual presentation can add to the understanding of what our mathematical model means and how it is related to real data, is in the analysis of covariance. Here the problem of simple regression is complicated by the fact that the pairs of correlated variables fall into groups between which there may be differences in mean values, in variability and in the slope of regression lines. If the main objective is to look for differences in mean values, we have to be clear that what we are looking for is not obscured by heterogeneity in variation or regression.

There are procedures which can be used to test these points, but when few observations are available such tests are very insensitive. There is now a good deal of information on the effects of departure from standard conditions, but this is not easy to put into precise form. When asked how he knows that his mathematical model is adequate, the practising statistician will perhaps answer that it is a matter of using judgment acquired through experience, by getting the feel of the data by a process which is partly intuitional. The beginner cannot have this confidence in his own judgment, but is there any surer way for him to acquire it than by making a practice of plotting his data and thinking round the picture in the light of what he knows of theoretical work on the robustness of statistical tests? Again, therefore, as in the preceding example, I shall venture to describe this process of "thinking round" in a little detail.

My illustration is based on the figures given in Table 2; these fall into four groups, each containing 10 pairs of correlated observations, x_{ti}, y_{ti} ($t = 1, 2, 3, 4$; $i = 1, 2, \ldots, 10$). They may be supposed to represent the weights (in lb.) of 40 animals, x before and y after a period of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

feeding with one of four special diets. The problem is to find out whether gain in weight is influenced by diet.

TABLE 2

Initial Weight (x) *and Final Weight* (y) *in lb. of* 40 *Animals*, 10 *of which were Fed on Each of Four Different Diets*

	Diet A_1		*Diet* A_2		*Diet* A_3		*Diet* A_4	
	x	y	x	y	x	y	x	y
	39	75	60	96	53	92	60	98
	48	74	56	91	50	91	53	94
	42	72	51	90	54	82	51	90
	45	70	53	86	48	80	48	90
	41	66	62	86	44	80	53	89
	37	64	46	80	47	75	58	87
	36	62	51	80	41	71	54	84
	41	59	54	80	36	70	44	82
	32	58	46	75	45	70	51	80
	39	58	44	72	32	64	40	77
Mean	40·0	65·8	52·3	83·6	45·0	77·5	51·2	87·1

The x, y points have been plotted in four separate charts, one for each diet in Fig. 5; on these charts we can as a first step usefully show as reference lines the overall mean values for the 40 observations $x.. = 47·1$, $y.. = 78·5$. Apart from the obvious correlation between x and y, we see at once that the initial mean weights, x, of the animals in the four groups differed considerably. This can be regarded as bad planning, but the difference being there can be allowed for.

The mathematical model to be used in the analysis of covariance supposes (i) that the regression of y on x within each group is linear, (ii) that the variation of y for given x about these regressions is normal; (iii) that for all four groups, and whatever the x value, the variance of y about the regression line is the same. If we draw in the four regression straight lines fitted to each of the sets of 10 observations (shown as broken lines in the diagram) with slopes

$$b_1 = 0·94,\ b_2 = 0·97,\ b_3 = 1·10 \text{ and } b_4 = 0·75$$

we have the data fully displayed for inspection. The general conclusion drawn from study of these pictures must I think be that when examining for differences in mean growth, there is no reason to doubt the usefulness of a model which involves the conditions (i), (ii) and (iii).* Further that it would be neither necessary nor profitable in this case to apply arithmetical tests of this assumption.

The diagram, however, might have shown a different state of affairs, one in which there was a strong suggestion of heterogeneity due to a few discrepant observations, or of departure from linearity in some or all of the group. It might then be difficult to decide how to proceed, but however tiresome this may be, if conclusions are to be true to the real data we must not allow all evidence of anomaly to be buried within a conventional arithmetical procedure.

Having drawn Fig. 5 in the first instance to check the mathematical model, it may next be used to give an added meaning to the analysis which we have now accepted as appropriate. We are considering possibilities that:

(a) The only difference between the groups lies in the x-values—the initial weights—so that the whole set of 40 points vary about a common regression line representing an average increase in weight during the period of observation, which is the same for all diets. This regression is estimated by the line with slope b_0 drawn through the common grand mean $x..$, $y..$ on each chart. (This is the line fitted to the 40 pooled observations.)

* As is frequently the case in problems of growth, we should anticipate that the variance of y would increase with x. Such an effect, if present in these data, does not appear to be sufficient to make nonsense of condition (iii).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(b) There is a differential effect of diet, so that the displacements of the mean points x_t, y_t (represented by open circles in each chart) are not consistent with chance fluctuation from a common line; but the slope of the regression of y on x is the same for all diets. This common slope is estimated by b_a.

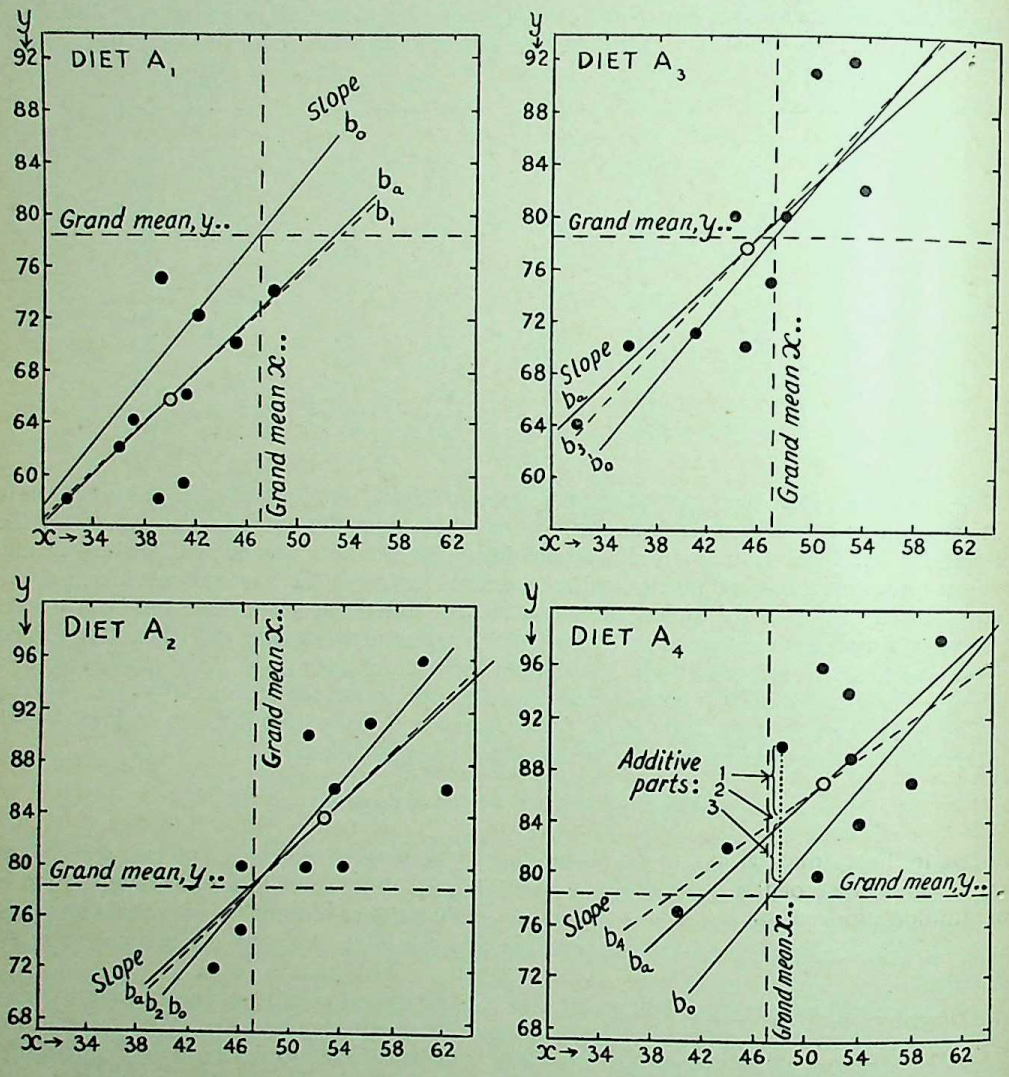


FIG. 5.—Analysis of co-variance.

(c) The rate of growth varies with diet so that we may have to consider four regression slopes, estimated by b_1, b_2, b_3 and b_4. To compare numerically with these four estimates quoted above, we have:

$$b_0 = 1{\cdot}22, \qquad b_a = 0{\cdot}95$$

With this picture in front of us, we can ask what exactly are the components whose sums of squares S_1, S_2 and S_3 are given in Table 3 below ? For one observation pair (x, y) in the fourth group of 10, the diagram shows the three component parts which have been squared and summed for all 40 observations. They are distances, parallel to the y axis, between—

(1) the observation point and the corresponding point on the line with slope $b_t (t = 1, 2, \ldots 4)$ drawn through the group mean;

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(2) the lines with slopes b_t and b_a drawn through the group mean;
(3) the line with slope b_a drawn through the group mean and that with slope b_0 drawn through the grand mean.

If we say that a comparison of the sums of squares of these components provides the appropriate means of distinguishing between the possibilities (a), (b) and (c) set out above, we are making an appeal to the imagination through the visual faculty. Mathematical theory has shown that we can make these comparisons most neatly by dividing each sum of squares by its appropriate degrees of freedom and referring to standard tables of the variance ratio F or of its logarithm. But the essential comparison appears geometrically in this diagram as the sum of squares of the three parts of a distance.

TABLE 3

Analysis of the Data in Table 2

Variation due to	*Sum of Squares*	*Degrees of Freedom*	*Mean Squares*
Deviations within each group from regressions b_t	$S_1 =$ 859·2	32	26·85
Differences among regressions b_t	$S_2 =$ 23·0	3	7·67
Deviations within groups from common regression b_a	$S_1 + S_2 =$ 882·2	35	25·21
Differences among corrected group means	$S_3 =$ 424·5	3	141·50
Total deviations about $y = y.. + b_0(x - x..)$	1306·7	38	

The arithmetical analysis summarized in Table 3 shows that $F = S_2/3 \div S_1/32 = 0{\cdot}286$ is not significantly large (nor exceptionally small) so that there is no ground for discarding the hypothesis that there is a common regression slope for all four diets. On the other hand $F = S_3/3 \div (S_1 + S_2)/35 = 5{\cdot}61$ is definitely significant, so that the displacement of the four group means from the common line

$$y = y.. + b_0(x - x..)$$

is clearly real. These conclusions obtained from the arithmetic area seen to be reasonable from the diagram.

What further question we ask about the established differences in effect of diet will depend on details of the experiment. If, as in the case of sheep trials for which data were given by R. A. Fisher (1947, Table 29.1), the treatments A_2 and A_3 represent the addition of two different constituents to the normal diet A_1, while A_4 represents the addition of both A_2 and A_3, we can examine for the significance of the two main effects and for the presence of a possible interaction. The diagram does, indeed, suggest that such additive main effects are present.

3. SEEING THE DATA AS A WHOLE

3.1. *Diagrams and Text-books*

I have described some of the ways in which I believe that visual aids may be of value in teaching the mathematical theory of statistics and in showing how its models are brought into relation with real data. I want now to add something about the importance of diagrams to the trained statistician as a means of helping him, in Hilbert's words, to keep in live *rapport* with his data.

The only published evidence of the way in which practising statisticians handle their data comes from the examples given in text-books or in papers dealing with the development of new methodology. Here it must be admitted that one finds but little evidence that statisticians examine geometrically what I shall term the pattern of their data before analysing them arithmetically. Certainly, publishers and editors are inclined to discourage frequent use of illustrations because of expense, but it is doubtful whether many of the writers of statistical text-books would find it easy to make full use of diagrams even if their publishers would allow them. The use of the visual imagination is rarely required by the mathematician and many of the authors of statistical texts have been primarily mathematicians.

In looking through these published numerical examples it is common to find instances where inspection by diagram would have enabled the author to add something of value to his comments

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and, sometimes, to avoid misinterpretation. These published examples are, of course, generally introduced to illustrate the working out of an arithmetical procedure based on the use of some mathematical model, and if the data used do not fit the model precisely the illustration of procedure may be none the less valuable. As mathematical statisticians we have no doubt most of us been concerned in illustrating new theory on data that are not altogether suitable. Recognizing this, I would disclaim any specific criticism of the writers whose data I shall now make use of. The examples were admirable for their purpose; I am using them for another purpose.

3.2. *Example* 1

My first example in this series is taken from *The Design and Analysis of Industrial Experiments* edited by O. L. Davies (1954, Example 4.4, pp. 124–30, 142–4) where it is used to illustrate methods of investigating the errors involved in sampling and analysing a compound material. The data reproduced in Table 4 are concerned with the chemical analysis of fertilizer for percentage of potash. Eight batches, each contained in a number of bags, were sampled by inserting an auger into a bag chosen at random from each batch. The sampling was repeated next day so as to give two independent samples, S_1 and S_2, from every batch. Each sample was then carefully divided into two sub-samples, the first sent to one laboratory to be analysed by method A the second to another laboratory for analysis by method R.

TABLE 4

Giving Percentage Potash Determined from Samples, by Two Methods

Sample	S_1		S_2	
Method	*A*	*R*	*A*	*R*
Batch 1	15·6	15·5	15·5	15·4
2	15·4	14·9	15·2	15·0
3	15·3	15·4	14·6	14·6
4	15·0	15·5	15·0	15·5
5	15·5	15·0	15·4	15·1
6	14·8	14·8	15·0	14·8
7	14·9	15·0	14·9	15·3
8	15·0	15·8	15·2	15·8

Let us now look at the plot of the data in Fig. 6. What first do we note? That with the exception of Batch 3 the difference between the two results for a given batch, using the *same* method of analysis (either A *or* R) is small; this difference can be ascribed to two sources

(i) Reproducibility in chemical analysis; the authors have reason to believe that this may be assumed to be the same for both methods of analysis, and they represent the variability by a standard error σ_0.

(ii) The error of sampling involved in taking two auger samples from each batch, and represented by σ_1.

But inspection of the diagram shows that there is clearly another type of error present. The determinations by methods A and R differ considerably; sometimes those for A are greater than those for R, sometimes the reverse (here again Batch 3 is exceptional). Clearly there is some bias which is not constant, and this variability in bias may be represented by a standard error, σ_2. To explain the phenomenon we have to appeal to information about the analytical processes. We are told that while A is a very accurate but lengthy method, R involves a much quicker technique but the results may vary considerably from batch to batch, perhaps being influenced by another constituent (other than potash) present in variable amounts in the batches of fertilizer.

Thus apart from batch-to-batch variation for which the diagram suggests some evidence, there are three standard errors to estimate σ_0, σ_1 and σ_2. If we omit Batch 3 from consideration a study of the pattern of points suggests at once (a) a large value for σ_2 (the spots are sometimes consistently well above, sometimes below the crosses), and (b) a value for σ_0 which is likely to be larger than for σ_1, as there is no consistent evidence that when one sample gives a higher reading than the other for method A it also does so for method R and vice versa. The result for Batch 3

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is, however, obviously exceptional since here a very large between-sample difference is consistently picked out by the two methods of analysis.

This immediately suggests that the statistical analysis should be carried out both with and without Batch 3. It is unnecessary to give the analysis of variance tables, but they lead to the estimates:

From 8 batches (as in Davies)

$$\hat{\sigma}_0 = 0{\cdot}098,\ \hat{\sigma}_1 = 0{\cdot}184,\ \hat{\sigma}_2 = 0{\cdot}378$$

Omitting Batch 3

$$\hat{\sigma}_0 = 0{\cdot}103,\ \hat{\sigma}_1 \text{ negligible},^*\ \hat{\sigma}_2 = 0{\cdot}410.$$

It is clear therefore that the figures for Batch 3 play a predominating part in determining the estimate of the error in sampling measured by σ_1. The preliminary examination by diagram shows

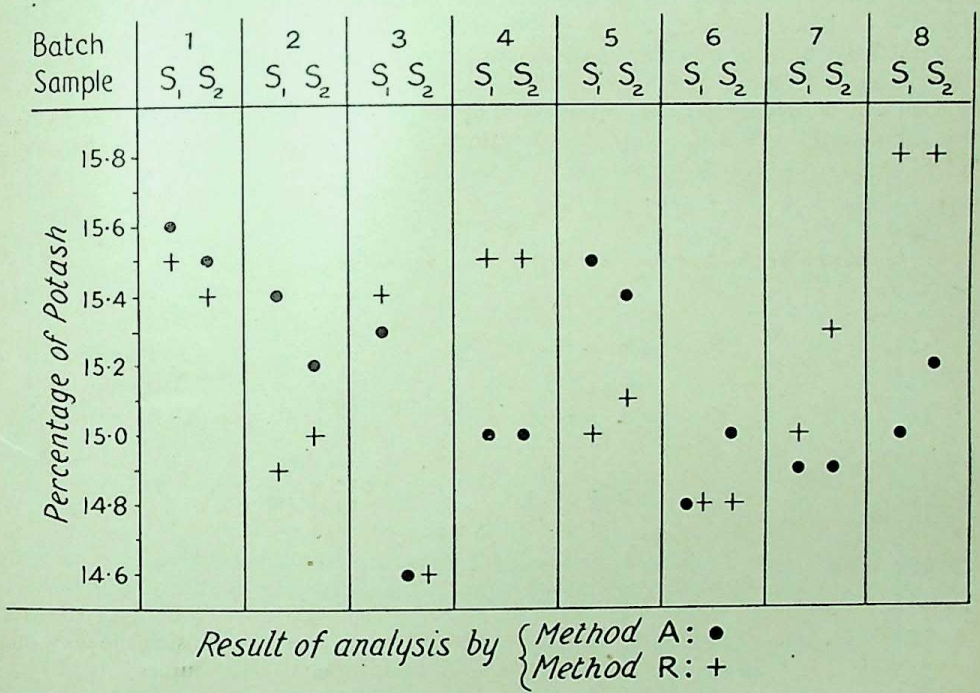


FIG. 6.—Errors of sampling and chemical analysis.

that this is likely to be the case before any arithmetical analysis is undertaken. What the implications of this anomaly are it would be for the statistician dealing with the investigation to decide.

3.3. *Example* 2

My next example is taken from a paper by C. R. Rao (1948, p. 62). It was used by this author to illustrate a method of determining an optimum "contrast" in multivariate analysis and for this purpose was entirely adequate to illustrate the technique and the type of conclusions which could be drawn from the technique's application. But it is instructive to examine the data rather more critically, as would be necessary if the multivariate theory were to be safely used in reaching useful experimental conclusions.

The data concerned the thickness of bark deposit on cork trees, measured by the weight of borings. It was to be expected that the bark would vary in depth according to the orientation on the tree and the problem set by theory was to find the linear function of the thicknesses in the four compass directions, North, South, East and West, which would pick out most decisively this lack of uniformity. The theory starts with three preliminary contrasting variables

* The estimate of $\sigma_0^2 + 2\sigma_1^2$ is less than that of σ_0^2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$y_1 = N + S - E - W, \qquad y_2 = S - W, \qquad y_3 = N - S$$

obtained from differencing the four original N, S, E and W measurements. When applied to the data it then proceeds on the assumption that the values y_1, y_2, y_3 as determined from measurements on 28 cork trees growing in a block of plantations may be regarded as a sample of three normally correlated variables.

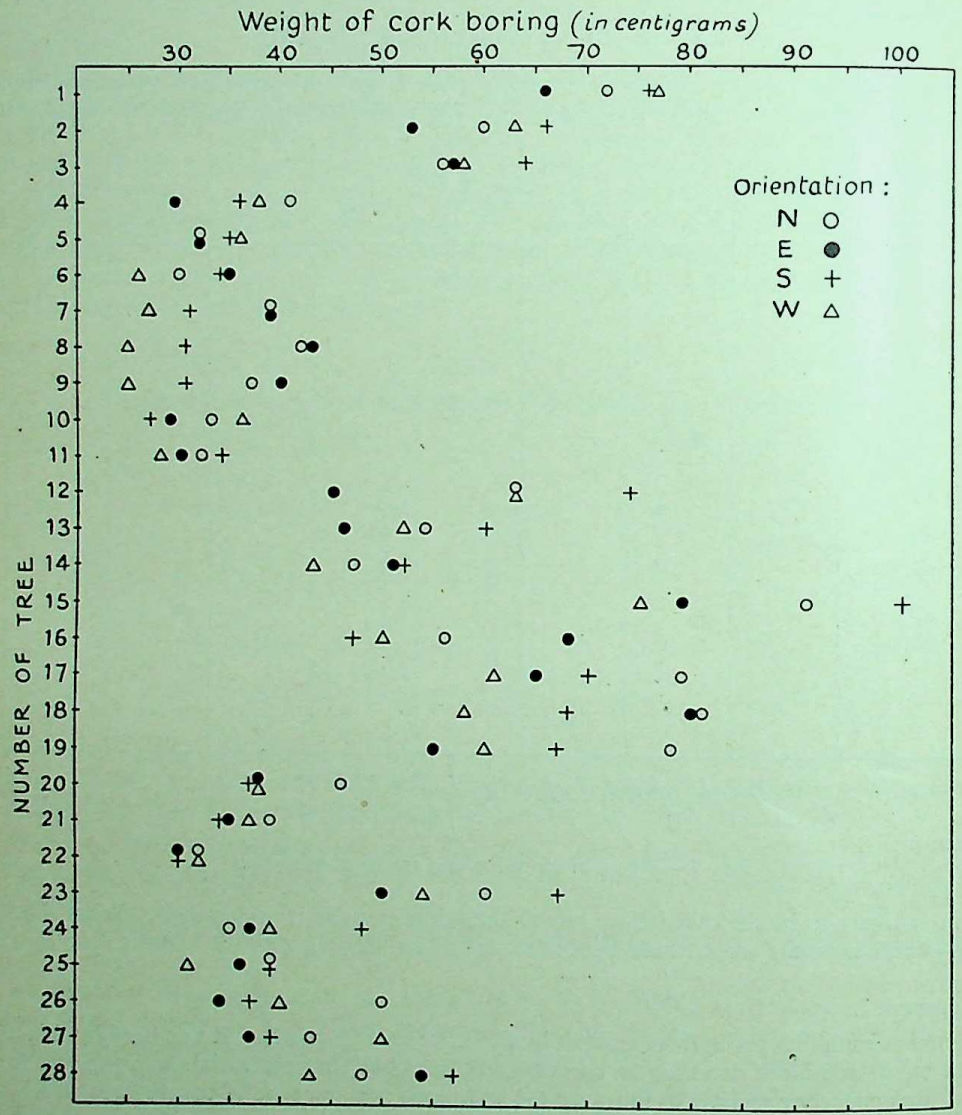


FIG. 7.—Measurements of bark on cork trees.

The original measurements in the four orientations have been plotted in Fig. 7. We see first of all the large differences between trees, but this main effect will be eliminated in the differences y_1, y_2 and y_3. We are not given any information about the order in which the 28 trees are numbered, but the eye is a sensitive instrument and it leaves us with the strong impression that, besides the very noticeable between-tree differences, there are differences in pattern (i.e. in the magnitude and sign of the y-values) which are associated with this ordering.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The question here raised by visual examination of pattern, is whether without further subdivision the data would be sufficiently homogeneous to justify useful conclusions being drawn from them by application of multivariate normal theory.* It would not be profitable to follow out this example further, but taken to this stage only it emphasizes the importance of some preliminary study of the statistical uniformity of data when several variables are used.

3.4. *Example* 3

A paper by W. L. Stevens (1948) provides another example which is relevant to the present discussion. He is concerned with developing an arithmetical procedure which may be applied systematically to test for any effects or interaction in experimental data collected under conditions where there has been a complete breakdown of the usual orthogonality restrictions. The data used are given in Table 5.

TABLE 5

Gains in Weight of Guinea-pigs on Different Diets (in Grams)

Type of Wheat in Diet	*Litter and Sex (M. or F.)*							
	I		*II*		*III*		*IV*	
	M.	*F.*	*M.*	*F.*	*M.*	*F.*	*M.*	*F*
A = soft (70%)	43	58	73	—	81	62	67	71
	58		59					
B = soft (100%)	93	60	75	71	101	76	100	—
	83		89					
C = hard (70%)	91	70	85	70	92	—	106	73
				58	88			
D = hard (100%)	89	—	98	69	105	—	109	76
	89			72	108			

They relate to an experiment (quoting from Stevens' paper) "designed to observe the effect on the growth of guinea-pigs of four diets distinguished by the four types of wheat which they included:

A = soft (70%), B = soft (100%), C = hard (70%), D = hard (100%).

The guinea-pigs were of two sexes and drawn from four litters, so there are three factors: sex, diet and litter." The experiment is therefore tri-factorial and, as the Table shows, presents a complete lack of orthogonality.

Stevens' paper is of considerable interest partly because of the warning which it gives of the severe penalties which may follow from neglecting the design of an experiment and partly from the workmanlike manner in which he tackles the problem, step by step. The arithmetical procedure involved is, however, formidable, involving the solution of many sets of simultaneous equations by a process of systematic adjustment set out in some ten pages of numerical tables! While the practising statistician will recognize the value of this paper as a contribution to the development and understanding of statistical methodology, he may I think legitimately ask: If I plot these 36 observations could I, or could I not, very quickly reach almost the same conclusions as come from this lengthy analysis,

(a) on tentative suggestions as to diet differences,
(b) on the essential need for more experiment, more carefully planned, if reliable information on the diet effects is to be obtained?

* It may be noted that if we apply Geary's (1936) criterion for kurtosis based on the ratio of the mean deviation to the standard deviation to this sample of 28 observations, we find that the distribution of y_1 is significantly leptokurtic at the 5 per cent. level and that of y_3 significantly platykurtic at the 1 per cent. level. These results also suggest possible lack of homogeneity.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The gains in weight from Table 5 have been plotted in Fig. 8. I wonder whether any trained statistician after studying the spot pattern would reach very different conclusions from those I did, noting down points as follows from my diagram before reading the conclusions which Stevens extracts from his analysis?

(i) There is a definite litter effect shown among the males, i.e. litters III and IV show greater increases in weight than I and II. It is unlikely that any litter × diet interaction could be established.

(ii) It is unlikely that a litter effect could be established among the females, although such a trend as appears resembles that for the males.

(iii) The most noticeable characteristic of the data is that, for the males, diet A is definitely inferior to the other three diets.

(iv) There is no similar effect which could be established for the females, although here, too, the average gain for diet A is in fact a little less than for B, C and D.

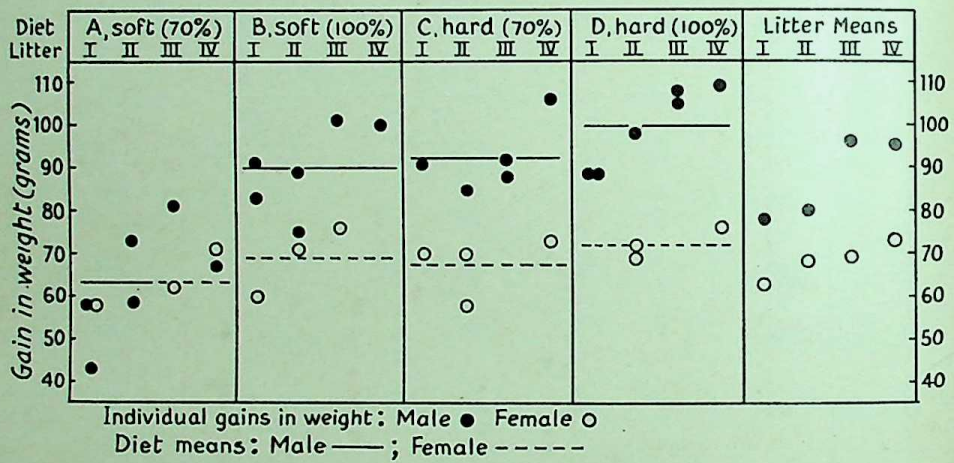


FIG. 8.—Gains in weight of guinea-pigs receiving different diets.

(v) As a result, the gain in weight for diet A appears to be the same for both sexes, but for the other three diets there is a marked difference between the males and females. Thus there seems to be a diet × sex interaction.

(vi) It is possible that there is a real advantage in "hardness" over "softness" in the wheat, and in "100%" over "70%"; the data are consistent with this, but it is the difference for males between A on the one hand and B, C, D on the other which overshadows all else.

It will be found that Stevens' analysis brings out just these points; the advantage of his treatment, of course, is that he can present his conclusions more precisely in terms of probabilities. For example, he finds that the diet × sex interaction is just significant at the 5 per cent. level. But in the matter of any positive contribution to the study of the effect of diet on guinea-pigs, he can only tell us that the results are inconclusive. Thus in connection with the diet × sex interaction he remarks that the observations are "consistent with either of the following hypotheses, indeed with many others:

"(a) That the difference between male and female is the same on diets B, C, D but different on A.
"(b) That the sex difference between male and female increases steadily with the response of either sex, i.e. the "better" the diet, the more the male will outstrip the female."

However we look at things, what is needed is further trial.

In making this comparison between the use of a diagram and the full arithmetical analysis of Stevens' data, I must emphasize that I would not normally suggest that visual inspection of a plot should be substituted for precise analysis. My contention is that a study of a diagram must always add something to our understanding of the results of analysis and that in cases where the analysis

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to be exact, must be very lengthy the diagram with certain simple exploratory calculations* may suffice for all practical purposes.

4. Concluding Remarks

Regarding geometry very broadly as a branch of science concerned with the presentation of both mathematical facts and observational material in a form in which their properties can be studied and illuminated with the aid of visual imagination, I have suggested that it has a part to play in various ways in the field of statistics. Geometry may first be useful in teaching; sometimes as an alternative to algebra as a means of establishing mathematical theory; sometimes in illustrating the relation between probability theory and random sampling; sometimes in making clear how a mathematical model fits a set of observational data.

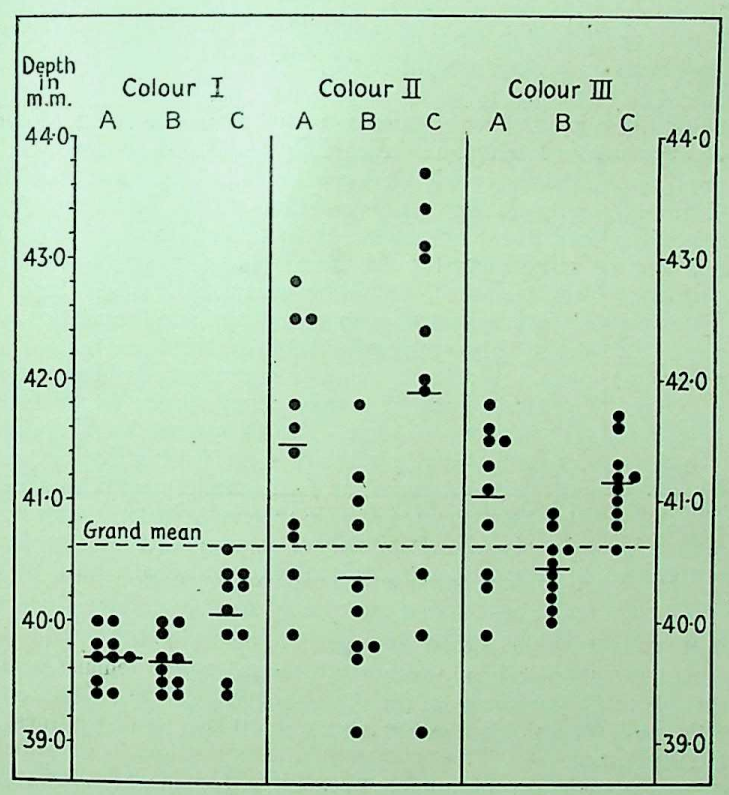


FIG. 9.—Matching of dye colours by three observers, A, B and C.

Beyond this I have considered how, in the practical interpretation of data, in the process of learning what the material means, examination of an appropriate diagram can supplement the usual arithmetical analysis, largely because only by this means is it easy to see the data as a whole. By suggesting that much can be learnt in this way, I run the risk of being accused of encouraging slapdash methods of handling statistical data, a step against the tradition of 60 years' development of statistics as a science. But there is also danger in the ease with which a standard text-book technique can be applied uncritically to observations to which the model does not fit. I am not suggesting that the visual imagination can replace hard calculation and the use of tables of probability functions. My thesis is that arithmetic and geometry are complementary.

The extremes to be avoided can be emphasized by considering two final diagrams. The first, Fig. 9, is a plot with mean values added of a part of more extensive data given a few years ago

* It will be noted from Table 5 that there are 9 pairs of animals, each member of which was of the same litter and sex and received the same diet. The 9 differences readily provide an estimate of error which can be used as a rough foot-rule in exploring the significance of possible effects.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in a University of London statistics practical examination. The figures are readings obtained from a colorimeter used for standardization purposes in the manufacture of industrial dyes. One cell of the instrument contains a fixed depth of a standard solution and the observer tries to make a colour match by adjusting the depth of solution in a second cell by means of a plunger. The diagram shows the adjusted depths in the second cell when the standard depth was 40 mm. In the figures here shown, each of three observers A, B and C made 10 matches with three colours. The candidates were asked to analyse and make a report on these data.

The diagram (which presents only about one-fifth of the total material provided) shows several things as obvious and others which it may be worth following up more carefully. Among the obvious things is the evidence that reproducibility in matching varies tremendously with the colour used. Although it should have been clear that no investigation of this kind could start before the observations had been looked at as a whole, the four or five students who attempted this question applied straight away an analysis of variance procedure which supposed that the standard deviation in each of the colour-observer groups was the same. These students had been urged more than once to look at their data before plunging into arithmetical analysis. If they cast common sense to the winds, was it from examination nervousness or the fault of a habit of mind, encouraged perhaps by school mathematics teaching and apparently confirmed by statistical text-books—an obstinate belief that diagram drawing is only an occupation for the mathematical beginner?

But if diagrams often give pointers of where to go in analysis and if they can sometimes establish results without need of arithmetical confirmation, it is unlikely that the most highly endowed statistical eye could ever be a substitute for the more objective probability measures based on calculation. Nevertheless there are some interesting questions involved here, concerning the sensitivity of the eye in picking out effects that are probably real and the relation of this sensitivity to conventional significance levels. My last diagram bears on these points.

Some little while ago I made a number of random samplings of different types and plotted the results for inspection. Fig. 10 contains six samples from a series in which five observations of a normal variable, y, fall into each of six arrays. All the samples were randomly drawn from a distribution in which there was a linear regression of y on x rising from left to right. I have drawn in the array means and the grand mean for each sample, quantities which can be easily calculated, but have not inserted the estimated sample regression lines. What conclusions could be drawn by inspection from any of these samples about the regression of y upon x?

It is likely that there would be general agreement among statisticians that:

(a) sample B 8 provides no justification for rejecting the hypothesis of independence;
(b) sample B 17 gives definite evidence of the existence of regression.

In these two instances, if a snap decision was needed, we should feel able to dispense with a test; but should we feel able to do this in any of the other four cases? A diagram of this type seems to bring out well the need for an objective test, and it could also be useful in teaching in getting across the idea of significance.

In fact, all six samples were drawn randomly from a distribution where there was linear regression of y upon x, measured by a slope of $\beta = 1{\cdot}31$, and on applying Student's test for the significance of the sample regressions, we have:

Sample . . .	B. 8	13	16	20	18	17
Slope, b . . .	0·34	0·73	0·91	1·22	1·15	1·89
t_0	0·69	1·38	1·63	1·81	2·76	4·55
$P\{t \geqslant t_0 \mid \beta = 0\}$* .	0·25	0·09	0·06	0·04	0·005	<0·001

* i.e., the probability that t could exceed the observed value, t_0, if the population regression line were horizontal with $\beta = 0$.

A comparison of this table with Fig. 10 leaves us with some interesting questions to turn over. Can we—indeed should we—ever base a conclusion on arithmetical analysis which we cannot at the same time see from inspection makes sense? Were a statistician to make the habit of relating to an appropriate diagram the probability level determined by every test of significance he applies, would he gradually acquire a more sensitive eye so that in time he could make a very shrewd guess from a diagram as to whether an effect were significant, let us say at about the 5 per cent. level? Would the same result apply to the confidence interval associated with an

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

estimate? Would this be a dangerous faculty to acquire, because it might lead to the abandonment of objectivity? Or would the gain in understanding outweigh any risk of misuse of the visual imagination? These are questions which I shall not attempt to answer, but I think they have an interest for us whether as teachers or practising statisticians.

The account which I have given has been necessarily to a large extent a personal one. I have taken a particular range of statistical problems which may well be regarded as somewhat limited, and have described how within this field the study of diagrams has helped me. Of this help I

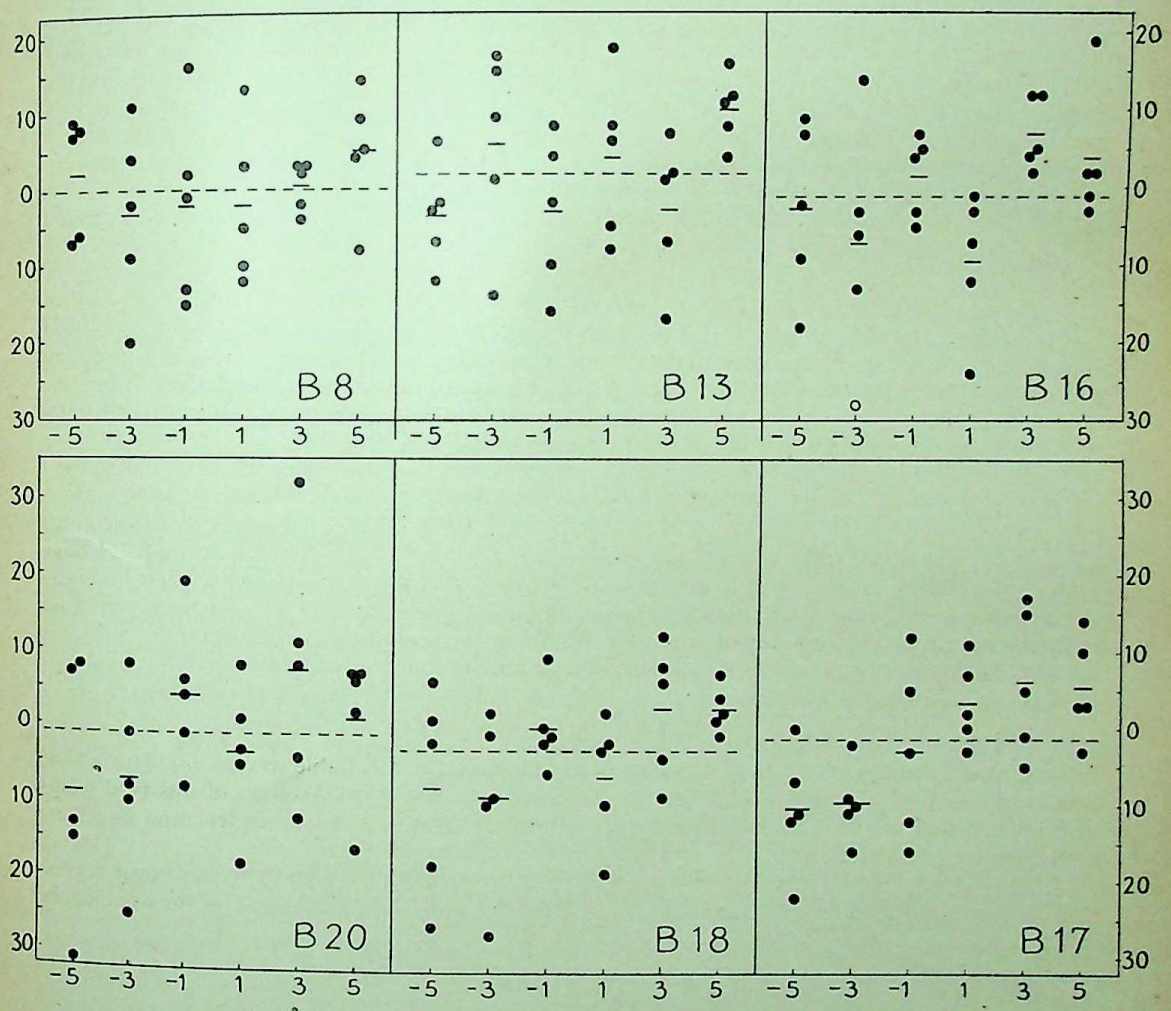


FIG. 10.—Six random samples of 30. Is the regression significant?

have no doubt whatsoever; as a teacher, help in seeing that a student's arithmetical results are probably right or certainly wrong; as an editor, help in understanding the purpose of an algebraic *tour de force*, by plotting the numerical data which the author has introduced to illustrate his results; help even without pencil and paper by seeing, through a kind of mental imagery, the character of mathematical problems to be solved. In general the mathematically-trained mind works no doubt in other ways, but the job of the average mathematical statistician is to learn from observational data with the help of mathematical tools. I am suggesting that in this process full understanding can only be obtained through visual aid and that if this is so, both in teaching and in writing of books, we need to give a good deal more consideration than we seem to do at

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

present to what was once described as the "geometry of statistics" and the "geometry of chance". This is the thought with which I should like to leave you.

References

DAVIES, O. L. (editor) (1947), *Statistical Methods in Research and Production*. Edinburgh : Oliver & Boyd.
—— (editor) (1954), *Design and Analysis of Industrial Experiments*. Edinburgh: Oliver & Boyd.
FISHER, R. A. (1915), "Distribution of the correlation coefficient", *Biometrika*, **10**, 507–21.
—— (1925), *Statistical Methods for Research Workers* (1st Edition). Edinburgh: Oliver & Boyd.
—— (1947), *The Design of Experiments*. 4th Ed. Edinburgh: Oliver & Boyd.
GEARY, R. C. (1936), "Moments of the ratio of the mean deviation to the standard deviation for normal samples", *Biometrika*, **28**, 295–303.
HILBERT, D. & COHN-VOSSEN, S. (1952), *Geometry and the Imagination*. New York: Chelsea Publishing Co. (translation of original *Anschauliche Geometrie* of 1932).
PEARSON, E. S. (1938), *Karl Pearson. An appreciation of some aspects of his life and work*. Cambridge University Press.
—— & HARTLEY, H. O. (1954), *Biometrika Tables for Statisticians*, **1**. Cambridge University Press.
RAO, C. R. (1948), "Tests of significance in multivariate analysis", *Biometrika*, **35**, 58–79.
STEVENS, W. L. (1948), "Statistical analysis of a non-orthogonal trifactorial experiment", *Biometrika*, **35**, 346–67.

PROCEEDINGS OF THE MEETING

Lord PIERCY (Immediate Past President). It is my privilege and a great pleasure to propose that the thanks of the Society be accorded to the President for his Inaugural Address. I will venture to say a word on that in a moment, but I should like at this point to offer my congratulations to him on being elected President of the Society. I enjoyed my term of office greatly and I feel sure he will be happy in his term. The Society, too, is to be congratulated; it is not very often that a son earns great distinction in a field which has already been adorned by a distinguished father, and Karl Pearson was one of the architects of mathematical statistics.

If I say something on the paper it is with great diffidence, because I do not pretend to be a mathematician. Two things struck me: and first, the interesting glimpse which the paper gave into the processes of Professor Pearson's own mind. I do not go as far as to suggest that there is anything unique even in Professor Pearson's powers of dealing in mental imagery. I believe—this is one of his own propositions—that visual apppreciation is a thing that can be greatly improved by training, even though there are many people of whom, at one time, I was one, who like to believe that they have no power at all of visualizing. Professor Pearson made an impressive plea for this approach—in teaching, and also as an aid to the mathematical statistician himself as a first approximation to understanding the pattern of the data, and in getting full benefit from the mathematical analysis. He seemed to me to make out a convincing case.

The "meat" in the paper, however, is contained in the mathematical examples, which you gentlemen will think over and digest, because the solid argument which Professor Pearson has brought forward in support of his thesis rests in those examples.

It occurs to me that one can draw from another field some support for Professor Pearson's general approach, that is the field of economic theory. In economics we are dealing with many quantitative matters, and with some observed regularities of behaviour; there are questions of proportion, of variation, co-variation and the rest of it; and it is difficult to resist the notion that in some sense or other thinking in economics is mathematical thinking. And sure enough at a certain date the professional mathematicians got down to it. Jevons, Marshall and Wicksteed are a few of the English names. Marshall and his school in particular by methods of graphical illustration made the mathematical aspects of economic reasoning intelligible to a race, which still subsists, of non-mathematical economists; and Marshall was able to build up a highly organised mathematical model of a self-balancing economic system, which the non-mathematical mind could grasp with great clarity. Marshall and the teachers whom he trained managed to give the non-mathematical economist not only understanding, but a genuine possibility of reaching out for further truth by the use of these methods. There is no doubt that the exposition of economic theory was radically changed by all this. I know that in my own person, since I was brought up on Mill and Cairnes.

If one continued this story it would be necessary to say that mathematical economics has become a large and highly cultivated field, a field in which to the non-mathematical mind there seem to be an enormous number of people who construct models on surprisingly simplified assumptions with the idea of exploring mathematically the properties of those models. So far good; but when an economist finds himself back in the practical problems from which he cannot be long detached, there is a danger that these abstract exercises may sub-consciously replace the effort to understand

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the facts themselves. That is the opposite of Professor Pearson's approach. He encourages us to start with the data and try to see the pattern of the data before we approach the question of the model; and it may be that in the modern development of mathematical statistics in the field of economics—what I suppose is meant by econometrics—we shall be able to use mathematical-statistical techniques to achieve results of practical validity which will supplement these and create a more Malthusian approach to these arbitrary theoretical structures.

However, I daresay that hidden somewhere in these remarks is a touch of a *Schadenfreude* which the non-mathematician feels for the mathematician; but I am certain that whatever our approach, mathematical or practical, we have all been interested in the President's paper and a great many of us will get even more out of it when we have the leisure to read it in detail.

Professor BRADFORD HILL (a Past President): I have first to lay before you an apology. A little more than a year ago, it was my privilege, as the penultimate president, to second the vote of thanks to Lord Piercy. Being at that time half way across the North Atlantic I was, unfortunately, unable to do so. However, in a brief written communication I did advise Fellows of the Society to note with great satisfaction that that would be positively the last occasion upon which the Society's Honorary Officers could call upon me to address them from this or any other platform. Alas, it is always dangerous to prophesy, particularly, as the Danish proverb says, about the future. I must apologise for making so foolish an extrapolation, for not observing that as Lord Heyworth most kindly and more than adequately acted as stand-in for me I might have to serve similarly but, very far from adequately, for Sir Ronald Fisher. I also begin to doubt my judgement of the powers of the Honorary Officers. I begin to wonder whether there is any escape from their orders until death do us part. This evening I make no promises. But in spite of having misled the Fellows of the Society it does give me very great pleasure to have this opportunity to support Lord Piercy in the proposal of a vote of thanks to the President. Slightly more than thirty years ago I sat at his feet at University College. He instructed me lucidly and patiently in the arts of quadrature and interpolation. So far does my memory run. It then jumps to a few years ago to up-state New York where, together with Dr. Joseph Berkson, I was present at a conference. In discussing the use of a particular function Dr. Berkson had written, "How should it be fitted? Everyone was using maximum likelihood, but instead I used least squares. . . . Why? I suppose partly it was natural cussedness, but also I think it was due to parsimoniousness. I had gone to statistical classes years before, paid tuition fees, and learned about least squares. Why should I waste all that money?" Now like Dr. Berkson I also went to statistical classes and learned many things. But this is where my advantage lay. As a disabled "hero" of World War I my tuition fees were paid for me by a grateful country. Without danger of psychological trauma I could cheerfully forget all these things. Until Dr. Berkson put the point so clearly I had not realized that this must be one of the very great assets of the welfare state.

Yet I would not have you think that I believe that that early instruction was wasted upon me, that I think meanly of it—merely because to-day I am hazy as to the nature of quadrature and regard osculatory interpolation as perhaps slightly immoral. Of the Cambridge philosophers Sidgwick and Jackson it was said that they taught younger men to seek for themselves and to seek with an exacting conscience. Nansen, on his return from the Antarctic, was asked whether many things happened that he had not foreseen. "No", he is said to have replied, "but many things that I had foreseen did *not* happen". Be very careful, said Pasteur, if you are looking for something—you may find it. In the sense of those quotations no one could pass through a course at the Galton Laboratory thirty to forty years ago and emerge unscathed. Taught by the President's distinguished father, by my predecessor Major Greenwood, by that most able teacher Miss Ethel Elderton, by Dr. J. O. Irwin, and by the President himself one was inevitably given a philosophy, a way of thinking, an approach to experiment and observation, in short an "exacting conscience" in planning, in analysis and in interpretation, that one would never lose however tenuous might be the hold of techniques and facts.

Coming to a later date it is just over twenty-one years ago that the Society founded its Industrial and Agricultural Research Section—now the Research Section. In this important development the President played a leading part. At the inaugural meeting there was, naturally enough, some discussion of the difficulties that the industrialist and the statistician would meet in understanding one another. "Dr. Pearson admitted", according to the Society's Centenary Annals, "that it was difficult for the statistician to express himself in simple terms, and added 'I know from personal experience that that is very hard work, but it has its fascination and it is essentially a problem of co-operation'." "If", the Annals continue, "the Section can promote a drive towards simplicity, there are others besides industrialists who will welcome their efforts". If Mr. Macrosty could see the last number of the Journal Series B, I rather doubt whether he would feel that the Section had

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

been notably successful in its drive towards simplicity. However that may be, the problem has clearly continued to fascinate Professor Pearson. In his address he has put before us one way of clarifying mathematical statistical concepts and of viewing numerical data.

Without transgressing our traditional rule that the president is infallible (on this occasion only, of course) I may, I hope, put forward some personal confessions. While many, perhaps the majority, of non-mathematical statisticians can perceive their data more simply in a diagram there are others, and I am one of them, who often become foolishly confused when confronted with them in that form. To take a specific example the President plots in Figure 6 the data set out in Table 4. "Let us now look", he says, "at the plot of the data in Figure 6. What first do we note?" I am myself so made that I first note some crosses and some black spots and I find great difficulty in going beyond that. But let me go back to the original data of Table 4 ; by inspection of the entries, with the aid of a few simple columns of differences, I have no difficulty in seeing the relative position of the two A's and the two R's, the contrast between the A's and the R's and the oddity of Batch 3. I find this a much easier task than with the diagram.

Similarly my first reaction to Figure 7 is to look for the original data in the hope that I can detect from them in their numerical form an interpretation which I find most difficult to determine from the diagram. In relation to Figure 8 the President wonders whether any trained statistician after studying the spot pattern would reach very different conclusions from those he did. I did not attempt it, but I reached (I think without "cheating") much the same conclusions by studying the original observations given in Table 5—again to me the easier approach. Clearly if the data were very numerous my "visual" assessment would fail me and some preliminary analyses would be essential. But that might be equally true, surely, of the diagram approach. Finally the President asks some questions. Perhaps I am on safer ground in endeavouring to answer them. Referring to his Figure 10 he says, "what conclusions could be drawn by inspection from any of these samples about the regression of y upon x?". I inspected these dots, though I again hankered after the data, and (without cheating) made the following decisions against which I set the actual significance levels given on p. 144:—

	B_8	B_{13}	B_{16}	B_{20}	B_{18}	B_{17}
Assessment by eye .	Not sig.	Not sig. ? border -line	Not sig.	Just sig.	Sig.	Highly sig.
Actual probability .	0·25	0·09	0·06	0·04	0·005	$<0·001$

Apparently, illogically, I have coped with the correlation diagrams though I fail hopelessly with the others. I can give no reason, but I offer the result as some answer to the President's query "were a statistician to make the habit of relating to an appropriate diagram the probability level determined by every test of significance he applies, would he gradually acquire a more sensitive eye so that in time he could make a very shrewd guess from a diagram as to whether an effect were significant, let us say at about the 5 per cent. level?" I never look, and have not looked for years, at a fourfold table, or any contingency or correlation table without assessing "significance" by eye—though not necessarily in terms of a probability. Though no doubt one remembers the successes and forgets the failures, I feel sure one can in course of time become very sensitive to the size of samples and to differences between them. I have no reason to suppose that that would not be equally true of those who can read diagrams, it might well be more so.

In thus exposing my personal frailty in relation to some forms of diagrams I have not intended a criticism of the President's enthusiasm. I am sure he is right in putting forward a form of simplification that may bring understanding to many students and subsequent wisdom to them in their handling of data in real life. I have merely been at pains to note the obvious, that human characteristics are so variable, that one man's meat is another man's poison, or, in other words, that there still is something to be said for the simple Robert Louis Stevenson philosophy that—

"The world is so full of a number of things
I'm sure we should all be as happy as kings."

It gives me great pleasure to second the vote of thanks to the President for his stimulating address.

The vote of thanks was put to the meeting and carried unanimously.

THE PRESIDENT: I began tonight by apologizing to you for delivering this address two months late and thanking you all for having come nevertheless in such strength to hear me. Now, my first

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

duty is to thank you very warmly for the vote of thanks you have given me and to say how much I have appreciated the friendly remarks about myself expressed by my predecessors in office.

Lord Piercy has raised the interesting question of the extent to which visual aids can be brought into play as a help in the understanding of econometric theory. I can only say that I think I should need these aids in attempting to master this theory myself. To Professor Bradford Hill I should like to say just two things. First, it is clear that the mantle of the old Industrial and Agricultural Research Section fell jointly upon the Research Section and the Industrial Applications Section, and it is to the Society's third journal, *Applied Statistics*, rather than to *Series B*, that we must look for a continuation of that drive towards simplicity in presentation of statistical concepts and methods which was in some of our minds twenty-one years ago.

In the second place I can assure Professor Hill that I found nothing critical in his expression of a strong preference for allowing his imagination to play around written numbers rather than geometric patterns. One of the interesting things in this subject is to find out how other people do set to work upon the analysis and interpretation of data. It almost seems that the difference between us lies in this. While he obtains his survey of the data as a whole by passing his eye over a series of differences written as numbers on a piece of paper, I find it easier to see the same differencies as distances between points in two or sometimes three dimensions. If I am lucky to have something of the draughtsman's or sculptor's power of comparing and ordering distances by eye, he is gifted with what may be the more valuable intellectual asset, a more highly developed power of abstraction.

As a result of the ballot taken during the meeting, the candidates named below were elected Fellows of the Society:—

Alan Don Airth.
Victor Sydney Baker.
David Harold Blackwell.
Ardron Geoffrey Henry Clarebrough.
Josephine Elizabeth Coggins.
Charles Leslie Cullen.
Edward Michael Holroyd.
William Henry Kruskal.
John Lawrence Mills.
George Lloyd Owen.
Elizabeth Leonard Scott.
John Victor Smart.
Ryszard Syski.
Peter Cheow Khoo Tan.
P. Thionet.
Herbert Roy Touchin.
John Christopher Turner.
Joseph Frederick Turner.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Analysis of Business Concentration: A Statistical Approach*

By P. E. Hart and S. J. Prais

National Institute of Economic and Social Research, London

[Read before the Royal Statistical Society, February 15th, 1956, the President, Proffssor E. S. Pearson, C.B.E., in the Chair]

Contents

PAGE

1. Introduction 150

2. Methods of Measuring Business Concentration 152

3. Comparative Statics 154
   - 3.1. Comparisons for the United Kingdom, 1885–1950 154
   - 3.2. Symmetrical Lorenz curves 156
   - 3.3. The log-normal hypothesis 157
   - 3.4. The law of proportionate growth 161

4. The Dynamics of Business Concentration 161
   - 4.1. The mobility of firms in a competitive economy 161
   - 4.2. Transition matrices for the United Kingdom, 1885–1950 162
   - 4.3. The components of changes in business concentration 165
   - 4.4. The effects of amalgamations 168
   - 4.5. The growth of firms 169
   - 4.6. Growth and concentration 171

5. Conclusions 175

References 175

Appendices:
- A. The Collection of the Basic Statistics 176
- B. Some Tests of the Reliability of the Data 177
- C. Standardized Rates of Regression for a Constant Time Interval 179
- D. Some Companies with Substantial Changes in Market Valuation 180

1. Introduction

What are the forces governing business concentration and how have these changed in recent years? These questions—important from the point of view of both economic theory and economic policy—have recently stimulated considerable research in this country and in the United States. The studies yielded by this research have almost invariably found difficulty in summarizing the varied phenomena of business behaviour to give an over-all view of the situation. For this reason we have set ourselves a rather narrow task in this paper: we avoid entirely any discussion of the good and bad points of monopoly and their implications for policy; instead, we have chosen to study one measurable indicator which plays an important part in any discussion of changes in monopoly. This indicator is the concentration of business activity and the object of this paper is to assess its changes through time, to analyse those changes into certain fundamental components and to assess their relative importance.

Business concentration, whether for particular industries or the economy as a whole, is often measured with the help of concentration ratios depending on the share of the largest firms in any particular industry, as in the well-known paper by Leak and Maizels read to this Society in 1945. The approach adopted here is more general in form and follows that of Berle and Means (1932)

* We are indebted to James Jefferys for helpful discussion on the general problems considered in this paper; and to Miss V. Farries, Miss B. M. Swift and Mr. V. Madge for assistance in compiling the basic data. Its preparation has been made possible by a grant under the Conditional Aid scheme for the use of counterpart funds derived from United States economic aid.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

who measured the share of the largest 200 non-financial corporations in the income and assets of all non-financial corporations. An investigation of this kind requires a knowledge of the size distributions of all companies, but in order to obtain this information conveniently, and on a comparable basis for a sufficiently long period of time, it has been necessary to restrict the material scope of the investigation to those firms which have shares quoted on the London Stock Exchange. The *Official Year Books* and the *Register of Defunct and Other Companies* provide a substantial amount of information on this portion of the business sector which makes it amenable to statistical investigation. Of course, analysis of concentration within the quoted portion of the business sector is useful in itself since it includes by far the greater portion of the largest firms in the economy. But more important is the fact that the accessibility of information allows an investigation of the relative importance of the constituent parts of the changes in business concentration: once this task is accomplished it may be possible to turn from the quoted portion and examine in a more enlightened manner the changes in concentration and their causes within the larger sector of mining, manufacturing and distribution generally.

The present approach owes much to certain statistical writers, in particular to Gibrat (1931) whose work on the log-normal distribution of firm sizes is followed in section 3 of this paper, and to Galton, whose work on the regression of hereditary characteristics is formally analogous to the approach adopted in section 4 below. We have been particularly fortunate in that J. Aitchison and J. A. C. Brown, working at the Department of Applied Economics in Cambridge, have recently been studying the log-normal distribution and its application to economic data; they have kindly allowed us access to the manuscript of their forthcoming Monograph on this subject, and have been helpful in the discussion of a number of the statistical problems that have arisen.

The discussion proceeds in the following order. In the next section various measures of business concentration are surveyed and it is argued that the classical statistical measures based on the variance of the frequency distribution are in certain ways superior to the concentration ratios which are usually calculated in analyses of monopoly power. The third section is concerned with a preliminary inspection of the data and an assessment of their value in measuring the changes in business concentration in the United Kingdom in the past half century; its main purpose is to show how the log-normal hypothesis is helpful in interpreting some of the observed properties of the Lorenz curves and to examine the implications of log-normality. The changes in the concentration of industry are partly due to the appearance of new firms, the bankruptcy and disappearance of old firms, and amalgamations between firms, on the one hand; and, on the other hand, to the growth of existing firms by the ploughing-back of profits and the raising of new capital. The relative quantitative importance of these factors is examined in section 4. It will be shown that the growth of those firms which are alive both at the beginning and end of a given period of time can be represented by a simple probability scheme with parameters which are fundamental in the analysis of changes in business concentration. With the help of this scheme, the mobility of firms in the size distribution is shown to be intimately related to the changes in business concentration over time. The conclusions to be drawn from this study are presented in section 5 and certain of the more technical matters are treated in a number of appendices.

Before proceeding with the discussion, two remarks are necessary. First, on the period covered: the statistical analysis in this paper has been carried back to 1885, but it is clear that the companies quoted in 1885 and 1896 are less representative of all manufacturing firms (and even of large manufacturing firms) than are the companies quoted in 1939 and 1950, and the estimates of changes in business concentration in the early periods are correspondingly less reliable guides to changes in concentration within the manufacturing sector as a whole. Because of this, the estimates presented for these early years should be regarded primarily as experiments in technique. The study of changes in concentration in a constant set of companies in the periods 1885–96 and 1896–1907 is, perhaps, not open to these objections to the same extent.

Secondly, on terminology: demographic terms have been used throughout the paper; a company is said to be "born" when it receives a quotation on the Stock Exchange, it remains "alive" until it becomes bankrupt, is amalgamated with some other firm, or is for any reason removed from the Stock Exchange List, when it is said to "die". The fact that a company will usually be in operation before it is born, in the sense of this paper, and may be operating after its death, is the principal difficulty in the use of Stock Exchange material and, as will be seen, requires specific comment throughout the paper.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2. METHODS OF MEASURING BUSINESS CONCENTRATION

The various measures of business concentration that have been used in past studies may be divided into two types*: first, those which depend on the share of total business activity controlled by a certain *number* of the largest firms in the industry or the economy; and secondly, those which depend on the share of a certain *proportion* of the firms. The former measure provides what is known as an absolute concentration ratio relating to, say, n firms (for example, in the study by Leak and Maizels $n = 3$, and in that of Berle and Means $n = 200$); the latter measure provides a percentage concentration ratio based on, say, the largest p per cent. of firms. Both types of measure require a knowledge of the total output of the industry considered; in addition, the second type requires a knowledge of the total number of firms in the industry (so that the number forming the top p per cent. may be calculated), whereas this is superfluous information in comparisons based on the absolute concentration ratio. It is difficult to see how any entirely satisfactory judgement on changes in monopolistic tendencies in an industry can be made without a knowledge of the changes in the number of firms engaged in that industry, but perhaps those who advocate the use of an absolute ratio as an analytical tool would argue that this information should be considered separately.

A generalization of the percentage concentration ratio is given in a diagrammatic form by the Lorenz curve; this shows the proportion of total business activity controlled by any given proportion of the number of firms. If the curve is a straight line so that, for any x, the largest x per cent. of firms control x per cent. of the output, all the firms would be of equal size and industry may be said to be completely unconcentrated. In general, firms are not of equal size and the largest x per cent. of firms will control more than x per cent. of the activity; to summarize the distribution it is then convenient to use the area between the diagonal and the actual curve as a measure of concentration. This area expressed as a proportion of the total triangle beneath the diagonal is known as the Gini coefficient and is here denoted by g.

The advantage of an absolute concentration ratio, as compared with a percentage concentration ratio or a Gini coefficient, is that it gives information in a form that is more immediately suitable for assessing the likelihood of monopolistic practices. When the industry concerned is closely defined and consists of only a small number of firms its nature can thus be well described by giving the proportion in the hands of, say, the largest firm and the largest three firms.

In the present study, however, our interest is in the development over time of the size distribution for the economy as a whole or, at any rate, for a broad sector of the economy. The questions are therefore rather broader in content; for example, it may usefully be asked on the present approach whether the large firms today control a greater part of the resources of the economy in relation to the remaining firms than they used to. If they do, there is a *prima facie* case for saying that opportunities for monopolistic practices are increasing in individual industries. But perhaps of greater significance are the social and political consequences of any increase in the concentration of economic power; it was largely this motive that led to the many detailed investigations of the Temporary National Economic Committee in the United States.

The analysis of concentration within the whole economy would amount to a study of monopoly if inter-industry competition were so important that any firm could be regarded as competing with every other firm, irrespective of industry classification, for consumers' expenditure.† But in addition to its relevance to monopoly, and to the political problems associated with the concentration of economic power, the study of changes in concentration within the whole economy throws light on the type of problem discussed by Edwards (1949, p. 92) and Galbraith (1952) in which the fact of large size is of economic interest in its own right.

For a study of this broader nature a summary measure of concentration based on the detailed distribution of firm sizes, such as the Gini coefficient, seems preferable to a concentration ratio based on only one tail of the distribution. A possible intuitive interpretation of the Gini coefficient which is of interest in this connexion is as follows. Suppose the difference in the sizes of two firms provides a measure of the degree of "dominance" that the one may exert over the other's price or output policy; of course, there are many other factors that affect dominance in any real sense, but they must be ignored for this interpretation. Then a measure of dominance for an

* A more detailed account will be found in the papers by Blair (1951) and Rosenbluth (1955).

† As argued, for example, by Triffin (1941), pp. 88–9.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

industry as a whole may be found by taking the mean difference, irrespective of sign, between all possible pairs of firms. If x_i denotes the size of the ith firm, then the measure is given by

$$\Delta = \sum_i \sum_j | x_i - x_j | /N^2 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where N is the number of firms. This coefficient, also due to Gini, is directly related to the concentration coefficient determined from the Lorenz curve by the equation

$$g = \tfrac{1}{2} \Delta/\bar{x} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where $\bar{x}$ is the average size of firms; this equation applies whatever the distribution of the x_i may be.* The Gini coefficient can therefore be interpreted as a measure of average dominance in the group of firms considered.

The coefficients Δ and g are, of course, only measures of the dispersion of the sizes of firms and while they have a certain intuitive acceptability there seems no reason why other, and more conventional, measures of dispersion should not similarly be used as measures of business concentration. The measure adopted in this paper is that of "relative dispersion" calculated as the variance of the logarithms of the sizes of firms. This is independent of the unit of measurement, but in contrast to the arithmetic variance is not independent of the origin of measurement of the sizes of firms (that is, x_i may be replaced by kx_i, where k is the rate of conversion between the units of measurement, without changing the value of this summary measure; but if x_i is replaced by $x_i - k$ its value will be changed). It can only be applied to measure the dispersion of items that are necessarily positive.

The variance of the logarithms is a measure that has many advantages in analytic work. First, it is argued in the next section that the size distribution of firms is approximately log-normal; this measure therefore provides a basis for an approximate description of the whole distribution. Secondly, on the assumption of log-normality, the classical tests of statistical significance are immediately applicable to the measurement of changes in business concentration. Thirdly, the measure is based on all the information in the size distribution so that a change in any part of the distribution has its appropriate effect on the value of the measure. Fourthly, since the measure is based on a sum of squares it may be partitioned into components, as in equation (3) below, and so provides a basis for disentangling the various possible factors making for a change in concentration. Certain of these properties are possessed by other concentration measures; for example, Mendershausen (1946, pp. 164–7) has shown that g may be partitioned into components, but the formulae are rather more complicated.

If this measure of concentration shows a change between two dates it is generally necessary to consider changes in other factors before coming to any conclusion on changes in the state of competition. For example, it may happen that concentration has increased as shown by the logarithmic variance, the Gini coefficient, or a percentage concentration ratio, but that this has been accompanied by an increase in the number of firms in the industry. There may thus be opposing forces affecting the degree of competition, and it may well be difficult in such circumstances to summarize the changes in the situation in a single measure. The view occasionally put forward, for example by Blair (1950, p. 8), that these difficulties invalidate the use of Lorenz curves and associated measures is hardly tenable in its extreme form; for an increase in the number of firms may be accompanied by a rise or by a fall in the Lorenz curve depending on how the dispersion is affected. What has happened is a question of fact, and the facts may have to be assessed by reference to the change both in the number of firms and in their dispersion. It is certainly as true here as in other applied sciences that it may be misleading to rely exclusively on a single indicator as a summary of a complicated situation.

A remark on the use of the term "concentration" may not be out of place here. When the sizes of firms are widely dispersed about their mean size, the large firms will be relatively very large and control a large proportion of the total assets or income of the industry; business concentration is then said to be high, the standard deviation is large, the distribution of sizes is widely dispersed and *not* concentrated. On the other hand, when the distribution is concentrated about

* See Kendall (1943), p. 44 for the usual proof of (2); a geometric proof is given by Woytinsky (1943), p. 252.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the mean size business concentration will be low. To avoid confusion, the term "concentration" will be used in this paper to refer to the concentration of industry and not of firms' sizes about their mean size.

3. COMPARATIVE STATICS

The principal matters examined in this section are the size distributions and corresponding Lorenz curves of quoted companies in the United Kingdom for selected years during the period 1885–1950. The comparisons made are static in the sense that no attention is paid to changes in the ranks of particular firms between different dates; the distributions and Lorenz curves are simply compared in their gross forms.

3.1. *Comparisons for the United Kingdom*, 1885–1950

The sector which has been chosen for the measurement of business concentration consists of those companies engaged in mining, manufacturing or distribution which are quoted on the London Stock Exchange. The definition of "business" is not the same as that of company in the legal sense, but consists of a parent company together with all subsidiary companies in which it has more than 50 per cent. interest, that is, a business unit as in the study of Leak and Maizels (1945, p. 144). It is possible to measure the size of such business units by the amount of their issued capital, the nominal amount of their quoted capital, the market valuation of their quoted capital and, for recent years, by their total net and total fixed assets. In principle, other measures of size, such as employment, may be suitable, but since few business units publish figures of employment the computation of such size distributions has to be left to the appropriate census authorities. The measure which is most convenient for the purposes of comparison over a period of time sufficiently long for the detection of changes in business concentration is that of market valuation. Market valuation has been calculated as the product of the selling price of each type of share and the number of such shares; it gives the market's estimate of the value of the business unit in terms, presumably, of its capitalized future earning power. A number of calculations have been made comparing the market valuation of companies with other measures of their size which show satisfactory correlations; these are reported in Appendix B. Further details relating to the Stock Exchange classifications and the method of abstracting these statistics are given in Appendix A.

TABLE 1

Size Distribution of Quoted Business Units in the United Kingdom by Market Valuation 1885–1950

Class	*Upper Limit* £'000	1885	1896	1907	1924	1939 *Official*	1939 *Combined*	1950
A	3	1	—	—	2	8	19	—
B	7	1	3	1	6	7	20	—
C	15	3	4	9	10	4	48	3
D	31	3	4	9	17	21	67	37
E	62	7	21	51	64	34	158	68
F	125	13	37	105	114	69	239	204
G	250	12	59	119	122	100	336	322
H	500	8	49	110	125	135	276	465
K	1,000	7	31	78	102	157	220	423
L	2,000	5	24	42	66	97	149	249
M	4,000	—	13	31	38	94	84	184
N	8,000	—	3	9	35	50	43	76
O	16,000	—	2	5	16	37	30	40
P	32,000	—	—	1	5	16	15	22
Q	64,000	—	—	1	3	4	4	5
R	128,000	—	—	—	1	3	3	2
S	256,000	—	—	—	—	—	—	3
T	512,000	—	—	—	—	1	1	—
	Total	60	250	571	726	837	1,712	2,103

The combined frequency distribution of business units quoted in the Stock Exchange categories, Breweries and Distilleries, Commercial and Industrial, and Iron, Coal and Steel for six years in

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the period 1885–1950, are shown in Table 1. For 1939, distributions are given for both the Official list and the estimated combined Official and Supplementary lists. The size classes are arranged in geometric progression, with the upper limit of any interval twice its lower limit. If the frequencies are viewed as being plotted against the logarithms of the sizes of business units then the chosen size classes would, of course, be equal. By assuming that the frequencies are concentrated at the centres* of the class intervals, it is possible to construct a series of Lorenz curves showing the changes in concentration in this sector over the period of 65 years; these are shown in Fig. 1 (for 1939 only the combined Official and Supplementary distribution is shown). The figures have

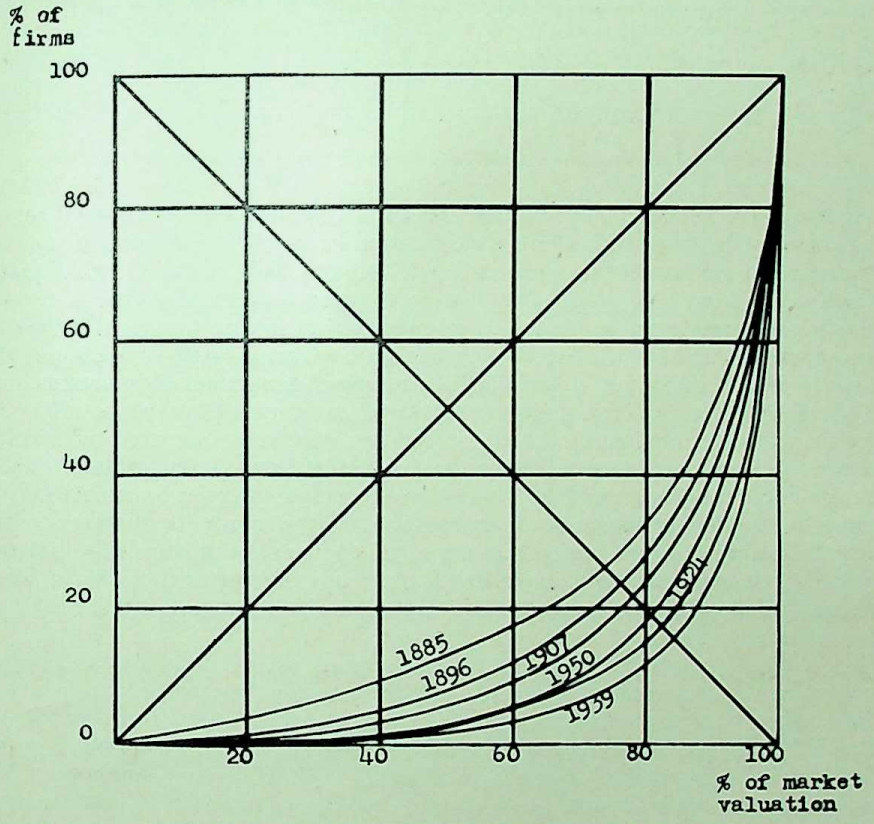


FIG. 1.—Lorenz curves for quoted business units in the U.K. 1885–1950, by market valuation

been cumulated from the largest business unit downwards, so that the traditional percentage concentration ratios can be taken from the graph immediately.

It will be seen that the Lorenz curves move away from the diagonal of zero concentration from 1885 to 1939, and there is a reversal for the post-war year 1950. With the possible exception of 1885, when the number of quoted companies was rather small, the curves exhibit a considerable degree of symmetry about the other diagonal. This symmetry of the Lorenz curves can be formulated as follows: if the top y per cent. of the units control x per cent. of the variate, then the bottom x per cent. of the units control approximately y per cent. of the variate.

The question which immediately springs to mind is: can the changes in concentration shown in Fig. 1 be taken to reflect changes that have actually taken place in the economy as a whole? The main difficulty in finding a satisfactory answer is ensuring that the sector of the economy

* The geometric mean was used for all except the end classes since an experimental study showed this to give a better approximation than the use of the arithmetic mean. For the top class the market valuations of the individual companies were used in calculating the Lorenz curves; for the bottom classes the arithmetic mean was used, as this was found to give a satisfactory approximation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

compared at the various dates is in fact comparable in coverage; from the totals given in Table 1, however, it is clear that the number of firms compared rose more rapidly than did the national income in real terms,* and the same is true of the market valuation of these firms.† It seems likely, therefore, that Fig. 1 reflects a substantially increasing sector of the economy and that no general conclusion can be drawn without further analysis.

If the actual share of the quoted firms in the economy as a whole were known, a more satisfactory appraisal of the situation would become possible; such estimates, using alternative measures of size, are being prepared at the Institute and will be reported in due course. Another method of appraising the trend in concentration is to compare a constant set of firms that are in operation at the beginning and end of a given period of time. Calculations of this kind find their place as part of the dynamic analysis of the next section; as will be seen from Fig. 5 below, these show movements of concentration similar in direction to those of Fig. 1 for each of the successive periods of time, though the magnitude of the movements is somewhat different.

3.2. *Symmetrical Lorenz Curves*

The symmetry of the Lorenz curves that has been observed here has presumably not been noted in earlier investigations of business concentration because of the tendency to examine only the upper portion of the distribution; only a few points are usually marked in published diagrams which are joined to form a polygon. However, in those instances where the full curve has been drawn the same symmetry may be observed; of published examples, we may cite the distribution of output among all manufacturing establishments in the United States in 1904 and 1939 in the recent study by Nutter (1951, p. 33), and the distribution of the output of ice-cream by establishments in the United States in the statistical text of Croxton and Cowden (p. 190). It is also interesting to note that the Lorenz curves for the two industries discussed by Blair‡ are both symmetrical. In the present investigations a large variety of data relating to the United Kingdom has been examined, ranging from the distribution of companies by gross true income assessed under Schedule D, to the distribution of establishment sizes by employment as reported in the Census of Production; they all suggest that a marked degree of symmetry is the rule rather than the exception. For reasons of space only one further set of examples is shown in Fig. 2; these relate to a number of industries in the United Kingdom in 1953–4 and illustrate the kind of variation that is found in practice.

In this case firm size has been measured by the net assets (that is, total assets less liabilities and provisions) as shown in the balance sheet of the company, taking the parent company's consolidated balance sheet where there are subsidiary companies. The companies examined are those operating principally in the United Kingdom (excluding finance companies) that were quoted on any United Kingdom Stock Exchange in 1953–4 and have a slightly larger coverage than those in Table 1.§

The curves shown in Fig. 2 relate (a) to all the 3,200 companies covered in the inquiry; (b) to the manufacture of clothing and footwear (excluding textiles), a relatively unconcentrated industry; (c) to chemicals (which includes pharmaceutical preparations and paint), and provides the most concentrated distribution; and (d) to the vehicles industry (which includes aircraft and cycles in addition to motor cars). Each of the three particular industrial groups consists of over a hundred companies and they differ considerably in their concentration coefficients||; the curves

* Prest's (1948) figures show an increase of approximately 50 per cent. in the real national income between 1907 and 1939.

† The aggregate market valuation of the firms as estimated from the frequency distributions rose approximately ninefold between 1907 and 1950 in current money units, and over threefold after correcting for rise in share prices (using the index of Smith and Horne (1934) linked with that published in the *L.C.E.S. Bulletin*).

‡ Blair (1950, Chart II, p. 8) gives Lorenz curves for the canned sea food industry and for the pharmaceutical preparations industry.

§ These companies are being examined in connexion with the National Institute's inquiry into the finance of public companies; we are obliged to Mr. A. Luboff and Mr. A. A. Sorrell for the information for Fig. 2. The numerical data on which this figure is based will shortly appear in a paper surveying some of the information yielded by the latter inquiry and its bearing on business concentration.

|| These examples of differences among concentration coefficients of different industries suggest that it may be worth investigating the extent to which changes in concentration for the whole of the industrial sector are the result of changes in the weights of the various constituent industries.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

are not perfectly symmetrical but sufficiently so as to require further investigation. The first curve relating to the net assets of all companies may be compared with the 1950 market valuation curve in Fig. 1. These curves are so close together that it has not proved practicable to show them both on the same diagram, thus supporting the view that with such very great differences in the sizes of companies it does not matter too much how size is measured.

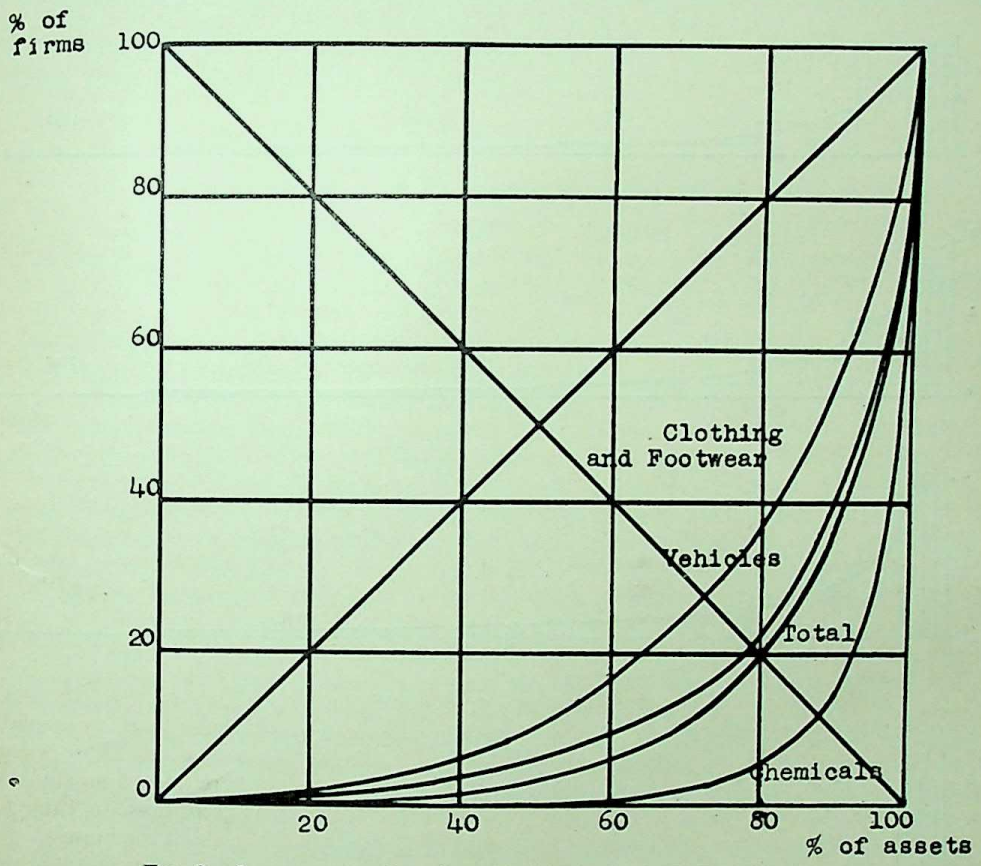


FIG. 2.—Lorenz curves for selected industries in the U.K. 1953–4 by net assets.

3.3. *The Log-normal Hypothesis*

The only frequency distribution known to us as yielding a symmetrical Lorenz curve is the log-normal distribution recently examined by Aitchison and Brown.* As a tool in the investigation of business concentration it has already been employed by Gibrat (1931) who devoted the third part of his book on *Les Inégalités Economiques* to the measurement of business concentration based mainly, but not exclusively, on French data. He argued that the log-normal distribution gave a good description of sizes of firms, and hence that the variance of the logarithms of the sizes—which is an estimate of one of the parameters of the distribution—provides an index of concentration.

In order to examine the fit of the log-normal distribution Fig. 3 has been prepared showing the distribution of the quoted companies in the U.K. for the selected years in the period 1885–1950 in a histogram form. In each case the theoretical two-parameter log-normal distribution has been drawn, the mean and variance of the distribution having been estimated from the corresponding

* The symmetry of the Lorenz curve is proved in their (1954) paper, p. 103.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

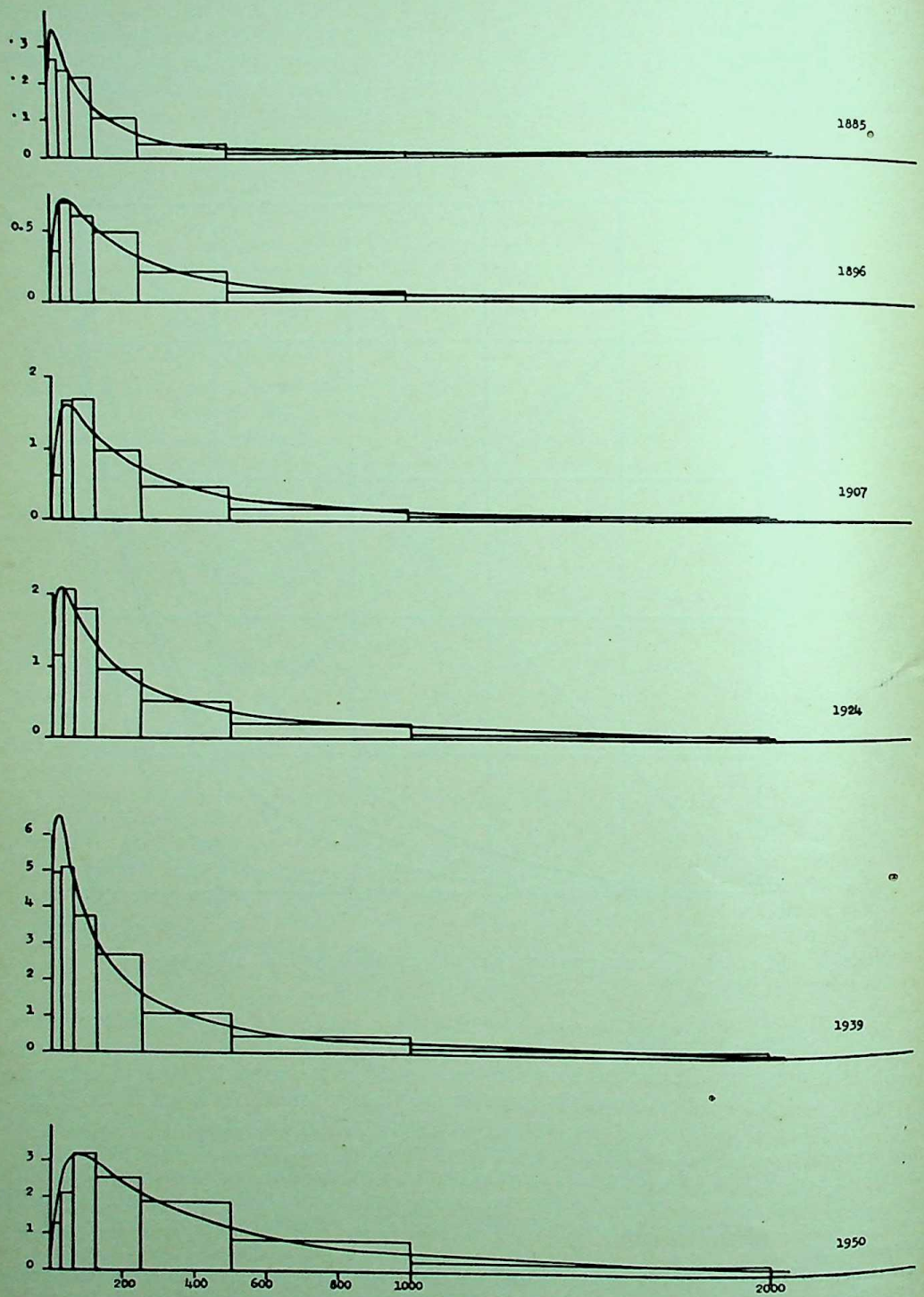


FIG. 3.—Frequency distributions of business units in the U.K. 1885–1950 with fitted log-normal curves

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

statistics of the transformed sample.* It has not been possible to include all the observations on the diagram on account of the large range they cover: a transformation to logarithms would have made this possible, but this did not seem desirable since it would not have shown the general shape of the distribution.

While the diagrams seem to support Gibrat's approach, it is desirable to provide some objective test of the satisfactoriness of the fit. To this end the third and fourth moments of the transformed distribution have been calculated together with their sampling errors, following the pattern of calculations set out by Fisher for testing the fit of a normal distribution.† The results are given in Table 2, in which the statistic g_1 provides a test of symmetry, and the statistic g_2 a test of kurtosis; both have an expected value of zero in normal samples.

TABLE 2

Tests of Normality of the Transformed Distributions of Fig. 3

	1885	1896	1907	1924	1939	1950
g_1	$-0{\cdot}41\pm0{\cdot}31$	$-0{\cdot}01\pm0{\cdot}15$	$0{\cdot}36\pm0{\cdot}10$	$0{\cdot}30\pm0{\cdot}09$	$0{\cdot}17\pm0{\cdot}08$	$0{\cdot}39\pm0{\cdot}07$
g_2	$0{\cdot}06\pm0{\cdot}61$	$0{\cdot}30\pm0{\cdot}31$	$0{\cdot}35\pm0{\cdot}20$	$0{\cdot}22\pm0{\cdot}18$	$0{\cdot}57\pm0{\cdot}16$	$0{\cdot}59\pm0{\cdot}14$

The table shows that the distributions tend to have positive values of both g_1 and g_2 indicating that the curves are somewhat skewed to the right and slightly leptokurtic. The discrepancies from normality, though small, are statistically significant from 1907 onwards, that is when the number of observations becomes large. Note, however, that between 1885 and 1950 the skewness shows a considerable net change (most of it having taken place by 1907), and that the kurtosis shows a fairly steady upward trend throughout. The reason for positive skewness is, perhaps, that the dividing line between quoted and unquoted companies becomes rather artificial at the lower end of the distribution; a number of the smaller companies that for most practical purposes are similar to quoted companies are thus not included in the frequency distributions examined here. In other words, quoted companies do not entirely form a naturally distinct set of similar objects and it may therefore be doubted whether much is to be learned (apart from gaining a better graduation of the data) from fitting a more involved form of curve. Attempts have, however, been made at improved fits: thus, in the case of sizes of firms (note that a business unit may consist of many firms) on which there is some Census of Production information, Tuck (1954) has suggested a modification of the log-normal hypothesis involving a large number of additional parameters (as many as eight are introduced, but some are based on extraneous information); and Rutherford (1955) has a modification involving two extra parameters which he applies to the income distributions of a number of countries.

If the simple log-normal hypothesis is accepted as a satisfactory approximation, the following properties would seem to be particularly relevant in analysing business concentration. First, there is the symmetry of the Lorenz curve which has been discussed above. Secondly, the Gini measure of concentration of a log-normal distribution depends only on the variance of the distribution (Aitchison and Brown, 1954, pp. 101–3); hence it is possible to derive the Gini measure from a calculation of the variance of the transformed distribution, and the latter is rather easier to handle in numerical work. Thirdly, given only one point on the Lorenz curve such as that usually provided by absolute concentration ratios it is possible to derive an estimate of the variance and the Gini coefficient. This is useful when comparing the available information on the concentration of industries either in the same country or, say, in two countries of differing sizes; it may then be important to judge whether a higher absolute concentration found in the smaller country

* No correction for grouping has been applied in any of the variance calculations in this paper so as to permit the straightforward application of equation (3) below. It is appreciated that there are more efficient methods of estimation, but in view of the fact that the quoted companies should perhaps be regarded as coming from the larger universe of all companies—so that the quoted sector comes from a truncated log-normal distribution—it seems doubtful whether there is any advantage to be gained at this stage from the pursuit of efficiency.

† See Fisher (8th ed. 1941), pp. 68–73. For the 1939 and 1950 distributions, which were based on a sample, the sampling errors have been calculated using the number in the sample as the number of degrees of freedom.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is due to the smaller number of firms or to a difference in the dispersions and this can only be done by making some assumption about the Lorenz curve.*

For the latter purpose it is convenient to use Fig. 4 which shows the Lorenz curves for log-normal distributions at intervals of 0·1 of the Gini coefficient together with the corresponding values of the variance of the logarithms (taken to the base 2). An example of its use is as follows: for the American blast furnaces industry in 1947 it is known (Kaplan, 1954, p. 93) that the largest four out of 33 firms, that is 12 per cent. of the firms, control 67 per cent. of the output; from Fig. 4 it is then seen that if the rest of the distribution is log-normal this corresponds to a Gini coefficient of about 0·75. A more convenient but equivalent method of calculating the Gini

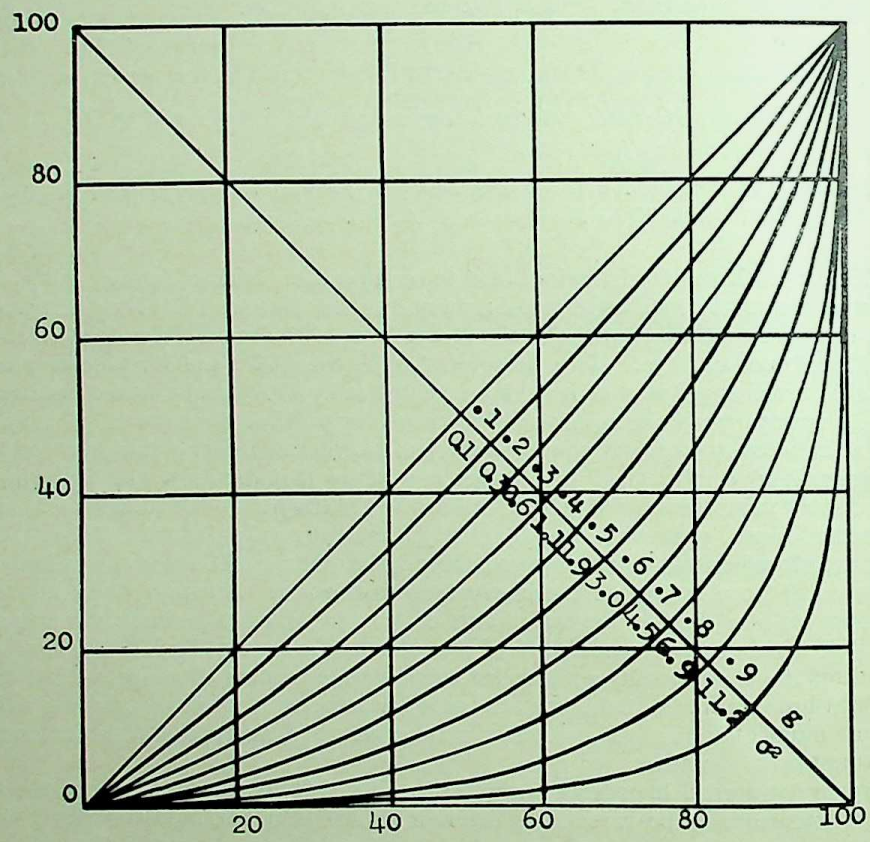


FIG. 4.—Theoretical Lorenz curves corresponding to log-normal distributions.

coefficient takes advantage of the simple relation between the log-normal distribution and its first moment distribution (Aitchison and Brown, 1954, p. 100, equation (5.4); the usual frequency distribution shows the proportion of firms having a certain employment, say; the first moment distribution would then show the proportion of employment in firms of that size). This is based on a specially constructed slide rule in which the probability scales cut from two pieces of probability graph paper slide against one another. For any setting of the scales, opposite readings on the scales give the points on the Lorenz curve of a log-normal distribution. It is a simple matter to graduate the positions so as to read off values of either the standard deviation of the distribution or the Gini coefficient. Unfortunately, the sampling errors of this method of estimation have not yet been worked out; nevertheless, in view of the wide range in which the absolute

* Cf. Rosenbluth (1954), pp. 335–6, for a problem of this type. He excludes the possibility of the dispersion of firms' sizes being greater in Canada than in the U.S.A., on the assumption that the points in his Table II (p. 341) lie on Lorenz curves which do not cross, and concludes that the higher absolute concentration ratio in Canadian industries are a result of their having fewer firms than the corresponding industries in U.S.A.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

concentration ratio may lie the method has been found useful and it is to be hoped that a formula for the variance will shortly be found.

Fourthly, on a more general level, the relation between the many alternative methods that have been proposed for measuring business concentration may be investigated with the help of the log-normal hypothesis; and as a result more direct, and perhaps more efficient, methods of analysis may be suggested. For example, the fact that a log-normal distribution and its first moment distribution have the same variance makes it possible to estimate the degree of concentration in the whole industry or economy even if, as is often the case, the total number of firms is not published or is not known. Thus, if the output of the largest firms and of the whole industry is known, it is possible to estimate the variance by plotting the upper points of the cumulative distribution of output among firms on logarithmic probability paper and this estimate of the variance, on the assumption that the distribution is log-normal, will be the same as that which would be derived by plotting the cumulative distribution of the numbers of firms.

3.4 *The Law of Proportionate Growth*

The discussion so far has been concerned with the static description of the size distribution of industry at a point of time. The next step is to discover how the distribution changes over time and, indeed, to understand the reasons for its adopting that particular form. The usual explanation of the origin of log-normal distributions is that they result from the action of a large number of small random factors acting upon a set of objects in a multiplicative manner; that is, the occurrence of event i to an object of size x would change it to size $\varepsilon_i x$, but would change an object of size y to $\varepsilon_i y$. This leads at once to a hypothesis of considerable economic interest, namely, that while a larger firm may have a better chance of increasing its size by a given absolute amount, the chance of a given proportionate increase is the same for firms of all sizes. This is the law of proportionate growth treated by Gibrat and is, perhaps, the most important consequence of the log-normal hypothesis. The available evidence on this law is examined below, but before turning to it, it is useful to consider certain dynamic aspects of business concentration in greater detail.

4. THE DYNAMICS OF BUSINESS CONCENTRATION

4.1. *The Mobility of Firms in a Competitive Economy*

If the firms in two industries, or two countries, have the same size distribution—and hence the same measure of concentration—it is still possible that their condition is very different from an economic point of view. For in one industry it may be that the largest firm at present was also the largest firm a century ago, and so on all along the size range; but in the other industry, or country, the state of competition and rate of innovation may be such that no firm holds its rank in the size distribution for very long. One may thus speak of industries that have a *rigid* structure and of those that have a *mobile* structure; industries that have both a high degree of concentration and a rigid structure are those in which one may suspect the existence of monopoly elements.

This line of thought leads to the examination of data which will answer certain dynamic questions. These ask, for example, whether the large firms have always been large and the small always small; whether the likelihood of bankruptcy or of absorption by another firm decreases as the firm grows in size; and whether new firms still always begin small and are only able to become larger with the lapse of time.

This shift in interest towards dynamic issues has a parallel in sociology which is worth noting. The earlier investigations into social classes and the structure of society were content to give the proportions of the population which consisted of landed gentry, tradesmen, paupers, etc.; but more recently the emphasis has changed to questions of social mobility such as, what proportion of the children now in class A had parents who were in class B; and what is the average number of generations spent by a family in a particular social class.* Naturally enough, certain of the

* Cf. the investigations in Glass (Ed. 1954) and Ginsberg (1931); it is, however rather difficult to draw a strict line between these investigations and the rather earlier studies of Galton which, as will be seen below, have a strong affinity with the present approach.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

formal methods that can be applied to the latter questions (Prais, 1955) can also be used in treating our problem of the growth of firms and their movement from one size class to another. The parallel between these two fields of study is, however, not merely formal; it arises in part from the substantive relation between the social and industrial hierarchies. Of course the study of business leadership—the importance of which has been emphasized in Schumpeter's writing—borders closely on both these fields, and the statistical approach is already being used in that field as well.*

4.2. *Transition Matrices for the United Kingdom*, 1885–1950

In order to present the information on the growth of firms in a concise manner it is convenient to divide the size scale into a number of ranges each of which forms a "size-class" and to show growth as movement from one such class to another. Provided that a sufficient number of firms is considered, there will then be little loss of information if the group is taken to be at the centre of each range. Tables of this sort can be regarded as giving information similar to that in the "league tables" showing the rank of a company in the size distribution at various dates (Kaplan, 1954, pp. 145–54). The differences are that the present tables are anonymous (they do not give the names of the business-units, which may be a disadvantage) and that the information is in a more summary form (which is an advantage for analytical purposes), so that a far greater number of companies can be handled conveniently. Two problems, in particular, have to be solved before such tables can be compiled.

The first concerns the choice of the end-points of the size-classes. On the basis of the conjecture (on p. 161 above) that all firms have the same chance of a given proportionate change in size it is to be expected that the frequencies will be conveniently distributed if the scale is divided into constant geometric size intervals. The intervals chosen were therefore the same as in the analysis of section 3.1 above (with the upper limit of the interval twice the lower limit) so that, for example, if a firm moves three classes upwards it will have increased its size approximately eightfold.

The second problem that requires consideration is the effect that changes in the price level during the interval of time considered may have on the size measure. If there were an unambiguous

TABLE 3

Changes in the Sizes of Business Units, 1885–96

Classes	A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	*Total*	*Births*	*Total 1896*
A	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
B	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	3
C	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	3	4
D	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	4	4
E	—	—	—	—	2	1	2	—	—	—	—	—	—	—	—	—	—	—	5	16	21
F	—	—	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	2	35	37
G	—	—	—	—	1	2	4	—	1	—	—	—	—	—	—	—	—	—	8	51	59
H	—	—	—	—	1	1	1	5	1	—	—	—	—	—	—	—	—	—	9	40	49
K	—	—	—	—	—	1	2	—	3	1	—	—	—	—	—	—	—	—	7	24	31
L	—	—	—	—	—	1	2	1	2	1	—	—	—	—	—	—	—	—	7	17	24
M	—	—	—	—	—	—	—	2	—	2	—	—	—	—	—	—	—	—	4	9	13
N	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	1	2	3
O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2
P	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Q	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
R	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
T	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total	—	—	—	1	5	8	11	8	7	5	—	—	—	—	—	—	—	—	45	205	250
Deaths	1	1	3	2	2	5	1	—	—	—	—	—	—	—	—	—	—	—	15	—	—
Total 1885	1	1	3	3	7	13	12	8	7	5	—	—	—	—	—	—	—	—	60	—	—

* Cf. the recent study by Copeman (1954) for the social origins of directors in the U.K.; and the work of Miss Erickson, at this Institute, based on a survey of certain sections of British industry covering well over a century.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

measure of the change in prices that could be used to deflate the end-points of the size-classes at the second date, it would clearly be desirable to do so. After such a process of deflation the upper limit of each class would still be twice the lower limit; its range would be changed in proportion to the index applied, but on a logarithmic scale its range would be unchanged and still equal to log 2. It is, however, rather difficult to find an index of prices suitable for deflating the sizes of firms, and this is true whether the size of a firm is measured by its income, the Stock Exchange valuation of its shares, or the net value of its assets as recorded in its accounts.

TABLE 4

Changes in the Sizes of Business Units, 1896–1907

Classes	A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	*Total*	*Births*	*Total* 1907
A	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
B	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1
C	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	2	7	9
D	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	1	8	9
E	—	—	—	1	4	1	5	1	—	—	—	—	—	—	—	—	—	—	12	39	51
F	—	—	—	—	3	13	12	1	—	—	—	—	—	—	—	—	—	—	29	76	105
G	—	—	—	—	1	5	15	8	2	1	—	—	—	—	—	—	—	—	32	87	119
H	—	—	—	—	1	4	13	17	6	1	—	—	—	—	—	—	—	—	42	68	110
K	—	—	—	—	—	—	6	14	10	4	1	—	—	—	—	—	—	—	35	43	78
L	—	—	—	—	1	—	—	—	6	13	3	—	—	—	—	—	—	—	23	19	42
M	—	—	—	—	—	—	1	2	5	3	5	1	—	—	—	—	—	—	17	14	31
N	—	—	—	—	—	—	—	—	—	—	2	1	—	—	—	—	—	—	3	6	9
O	—	—	—	—	—	—	—	—	—	—	1	1	1	—	—	—	—	—	3	2	5
P	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1
Q	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	—	1
R	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
T	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total Deaths	—	—	—	1	12	24	52	43	29	22	12	3	2	—	—	—	—	—	200	371	571
	—	3	4	3	9	13	7	6	2	2	1	—	—	—	—	—	—	—	50	—	—
Total 1896	—	3	4	4	21	37	59	49	31	24	13	3	2	—	—	—	—	—	250	—	—

TABLE 5

Changes in the Sizes of Business Units, 1907–24

Classes	A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	*Total*	*Births*	*Total* 1924
A	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	2
B	—	—	1	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	3	3	6
C	—	1	—	1	3	3	—	—	—	—	—	—	—	—	—	—	—	—	8	2	10
D	—	—	—	—	3	5	3	—	—	—	—	—	—	—	—	—	—	—	11	6	17
E	—	—	—	2	14	16	2	1	—	—	—	—	—	—	—	—	—	—	35	29	64
F	—	—	—	—	9	33	34	2	1	—	—	—	—	—	—	—	—	—	79	35	114
G	—	—	—	—	2	13	30	23	4	1	—	—	—	—	—	—	—	—	73	49	122
H	—	—	—	—	3	3	14	36	14	1	—	—	—	—	—	—	—	—	71	54	125
K	—	—	—	—	1	2	9	20	26	5	2	—	—	—	—	—	—	—	65	37	102
L	—	—	—	—	—	1	3	5	13	13	5	—	—	—	—	—	—	—	40	26	66
M	—	—	—	—	—	—	1	3	4	11	8	—	—	—	—	—	—	—	27	10	38
N	—	—	—	—	—	—	—	1	3	5	6	3	—	—	—	—	—	—	18	17	35
O	—	—	—	—	—	—	—	2	—	1	6	4	2	—	—	—	—	—	15	1	16
P	—	—	—	—	—	—	—	—	—	—	—	2	3	—	—	—	—	—	5	—	5
Q	—	—	—	—	—	—	—	1	—	—	1	—	—	—	1	—	—	—	3	—	3
R	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1	—	1
S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
T	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total Deaths	—	1	2	4	36	77	96	94	65	37	28	9	5	1	1	—	—	—	456	270	726
	—	—	7	5	15	28	23	16	13	5	1	—	—	—	—	—	—	—	113	—	—
Total 1907	—	1	9	9	51	105	119	110	78	42	29	9	5	1	1	—	—	—	569	—	—

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 6

Changes in the Sizes of Business Units, 1924–39

Classes	A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	*Total*	*Births*	*Total* 1939
A	—	—	—	1	3	1	2	—	—	—	—	—	—	—	—	—	—	—	7	1	8
B	—	—	—	1	4	1	—	—	—	—	—	—	—	—	—	—	—	—	6	1	7
C	—	—	—	—	1	2	—	—	—	—	—	—	—	—	—	—	—	—	3	1	4
D	—	—	—	—	6	4	5	1	—	—	—	—	—	—	—	—	—	—	16	5	21
E	—	1	3	1	8	7	3	3	—	—	—	—	—	—	—	—	—	—	26	8	34
F	—	—	—	2	13	21	9	11	2	1	—	—	—	—	—	—	—	—	59	10	69
G	—	—	—	1	4	24	19	12	5	3	—	—	—	—	—	—	—	—	68	32	100
H	1	—	—	—	2	3	21	21	13	3	—	1	—	—	—	—	—	—	65	70	135
K	—	—	—	—	—	5	9	25	31	11	6	1	1	—	—	—	—	—	89	68	157
L	—	—	—	—	1	2	4	10	15	14	4	1	—	—	—	—	—	—	51	46	97
M	—	—	—	—	2	4	1	9	9	13	8	9	2	—	—	—	—	—	57	37	94
N	—	—	—	—	—	—	1	—	1	6	8	9	2	—	—	—	—	—	27	23	50
O	—	—	—	—	—	—	1	—	2	3	4	9	4	—	—	—	—	—	23	14	37
P	—	—	—	—	—	—	—	—	—	2	1	2	5	2	—	—	—	—	12	4	16
Q	—	—	—	—	—	—	—	—	—	—	—	—	1	1	2	—	—	—	4	—	4
R	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	2	1	3
S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
T	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1	—	1
Total	1	1	3	6	44	74	75	92	78	56	31	32	15	4	3	1	—	—	516	321	837
Deaths	1	5	7	11	20	39	44	32	22	10	7	3	1	1	—	—	—	—	203	—	—
Total 1924	2	6	10	17	64	113	119	124	100	66	38	35	16	5	3	1	—	—	719	—	—

TABLE 7

Changes in the Sizes of Business Units, 1939–50

Classes	A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	*Total*	*Births*	*Total* 1950
A	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
B	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
C	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	2	1	3
D	6	3	9	—	3	9	—	—	—	3	—	—	—	—	—	—	—	—	33	4	37
E	3	—	6	9	18	8	3	—	—	—	—	—	—	—	—	—	—	—	47	21	68
F	3	3	10	21	42	46	25	1	1	—	—	—	—	—	—	—	—	—	152	52	204
G	—	3	9	12	33	62	64	3	7	3	1	—	—	—	—	—	—	—	197	125	322
H	—	—	—	9	31	61	124	69	8	3	—	—	—	—	—	—	—	—	305	160	465
K	3	3	—	3	—	15	61	114	73	8	—	—	—	—	—	—	—	—	280	143	423
L	—	—	—	—	—	1	12	39	86	49	1	—	1	—	—	—	—	—	189	60	249
M	—	—	—	—	—	1	6	10	26	68	44	3	1	—	—	—	—	—	159	25	184
N	—	—	—	—	—	—	—	—	6	7	35	19	2	—	—	—	—	—	69	7	76
O	—	—	—	—	—	—	—	—	—	—	1	20	17	—	—	—	—	—	38	2	40
P	—	—	—	—	—	—	—	—	—	—	—	—	8	13	—	—	—	—	21	1	22
Q	—	—	—	—	—	—	—	—	—	—	—	—	—	2	3	—	—	—	5	—	5
R	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	2	—	2
S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	—	1	3	—	3
T	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Total	15	12	34	54	127	204	296	236	207	141	82	42	29	15	4	3	—	1	1502	601	2103
Deaths	4	8	14	13	31	35	40	40	13	8	2	1	1	—	—	—	—	—	210	—	—
Total 1939	19	20	48	67	158	239	336	276	220	149	84	43	30	15	4	3	—	1	1712	—	—

Fortunately the problem can, for the purposes of the present analysis, be avoided on the assumption that a change in prices would affect all firms in the same proportion; for, if prices have doubled between the two dates, all firms would appear to have moved up by exactly one size class too many (given that the size classes are defined as in the previous paragraph) and a re-labelling of the classes at the second date is all that is required to convert the table to real terms. If prices have changed by a factor that is not an exact power of two, the sole further complication is that some interpolation may be required. The data can therefore be compiled on the basis of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

current monetary units, and the question of conversion into real terms dealt with when required.

On this basis Tables 3 to 7 have been compiled showing how the business units quoted on the London Stock Exchange have changed in size between the years 1885, 1896, 1907, 1939 and 1950. The size classes are designated by the code letters given in Table 1 above. The penultimate column in each table, headed "births", gives the number of firms alive at the second date which could not be traced at the first date. For the greater part, no doubt, these will relate to companies that have received a quotation during the period and will have been in operation at the base date as unquoted companies. The penultimate row similarly relates to "deaths" and will include, among others, companies that were independent at the base date but are subsidiaries of other companies at the second date. When an amalgamation occurred within the quoted sector the company with the largest market valuation before the amalgamation was regarded as acquiring the other members of the amalgamation, so that the largest company was promoted while the others were regarded as deaths.*

If an element in any of these tables is divided by the sum of the elements of the column in which it stands, the resulting fraction can be regarded as giving the probability of a firm's moving from one size-class to another; in particular, the ith element in the jth column (denoted in the footnote below as p_{ij}) will give the probability of a firm's moving in that time period from class j into class i. Such an array of probability elements is termed a "transition matrix"; its columns are standardized and may be compared with one another in order to assess how growth affects firms of different sizes.†

4.3. *The Components of Changes in Business Concentration*

Between any two dates business concentration may change because of (i) the entry of new firms, (ii) the exit of old firms (including their amalgamation with other firms), and (iii) changes in the size distributions of those firms, that are in existence at both dates. In this section an attempt is made to assess the relative magnitude of these components and, in the following section, the effects of amalgamations are assessed in a similar manner. It will be seen that the past emphasis on the effects of amalgamations is somewhat misplaced: amalgamations here emerge as only one of many contributing factors in a complex situation.

The simplicity of the present analysis depends on the use of the variance as an indicator of concentration. If two collections of objects are combined with means $\bar{x}_1$ and $\bar{x}_2$ and variances of σ_1^2 and σ_2^2, then the variance of the combination is given by

$$\sigma^2 = w_1\sigma_1^2 + w_2\sigma_2^2 + w_1w_2(\bar{x}_1 - \bar{x}_2)^2 \quad . \quad . \quad . \quad . \quad (3)$$

where w_1 and w_2 are the proportions of the objects in the two groups. Thus we have as particular cases of this equation, (a) if the two collections have the same mean and variance, the combination will also have the same variance; (b) if they have the same mean but differing variances, the variance of the combination will be a weighted average of the component variances; (c) if they have the same variance but differing means, the variance of the combination is always higher.

Table 8 summarizes the figures given in the basic tables of the previous section which are required in calculations of this sort. The means and variances of the various groups of companies have been calculated from the latter tables taking class G as having a zero working mean, and treating each size-class as having unit length; the results are therefore in terms of logarithms to

* The relatively large number of deaths recorded in these tables led us to make a more intensive inquiry into the lethal factors at work, with the help of the *Register of Defunct and Other Companies*. On the basis of a sample of 400 deaths examined for the period 1885–1950, it seems that some 37 per cent. of the deaths are due to liquidation, 34 per cent. to amalgamation and 27 per cent. to the loss of active quotation. Liquidated companies include both those that have been wound up as failures and those that have been voluntarily liquidated, possibly even while operating on a profitable basis, prior to the acquisition of their assets by other firms.

† On the basis of such a matrix it is possible to develop a theory of the changes in the size distribution of firms, as in the social mobility problem. Thus, if there are $s_j(t)$ firms in the jth class at time t, then $s_i(t+1) = \sum_j p_{ij}\, s_j(t)$. Neglecting births and deaths, and on the assumption of constant probabilities, it is possible to construct a series of Lorenz curves showing the expected changes in business concentration. In practice this development is of limited use on account of the necessity, and difficulty, of making realistic assumptions about the rates of births and deaths.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 8

Summary of Changes in the Numbers, Means and Variances of Quoted Companies in the United Kingdom, 1885–1950

	1885–1896			1896–1907			1907–1924			1924–1939			Units: Log_2 1939–1950		
	N	$\bar{x}$	σ^2	N	$\bar{x}$	σ^2	N	$\bar{x}$	σ^2	N	$\bar{x}$	σ^2	N	$\bar{x}$	σ^2
1. Total at first year . .	60	−0·37	4·17	250	0·56	3·85	569	0·57	3·79	719	0·89	5·61	1712	0·53	6·01
2. Deaths in period . .	15	−2·53	2·92	50	−1·02	4·22	113	−0·35	3·18	203	−0·06	4·73	210	−0·70	4·75
3. Survivors at first year .	45	0·36	2·50	200	0·96	2·97	456	0·80	3·68	516	1·26	5·45	1502	0·71	5·95
4. Survivors at second year .	45	0·98	4·64	200	1·13	3·85	456	0·96	6·16	516	1·56	7·66	1502	1·68	4·66
5. Births in period . .	205	0·47	3·63	371	0·28	3·55	270	0·77	4·54	321	2·12	4·62	601	1·11	2·34
6. Total at second year .	250	0·56	3·85	571	0·58	3·82	726	0·89	5·56	837	1·77	6·57	2103	1·52	4·07

the base 2, taken as deviations from an average company with a size of £177,000 (that is, the geometric mean of £125,000 and £250,000 which are the end-points of that class-interval).

As an example, we may take the period 1907–1924 during which the variance increased from 3·79 to 5·56. The companies that died during the period had a variance similar to those in existence at its beginning, but their mean size was rather smaller; in fact, the geometric mean size of the deaths was only 53 per cent. [$\text{antilog}_2 (-0{\cdot}35 - 0{\cdot}57) = \text{antilog}_2 - 0{\cdot}92$] of the total alive in 1907. As a result, the variance in 1907 of those alive throughout the period was reduced to 3·68. The next item in the table shows that in consequence of the relative changes in the sizes of surviving companies during the period, the variance increased to 6·16; this is unquestionably the largest factor in this period. It is slightly modified by the births which have both a smaller mean, by 12 per cent., and a smaller variance, by 26 per cent.; and this reduces the observed variance from 6·16 to 5·56.

A study of this and the previous tables leads to the following conclusions: (a) the business units that cease to exist in any period of time are smaller, by about a half, than the average size of units alive at the beginning of the period. This is shown by the difference of approximately a unit between the mean sizes in rows 1 and 2 of Table 8. The variance about that mean size is also generally smaller. Further, an examination of the basic tables shows that the probability of death decreases as the company grows in size. For example, from Table 7 it is seen that of the 96 companies in the top size-classes N–T that were alive in 1939, only 2 died; whereas of the 87 smallest companies in classes A–C, 26, that is 30 per cent., had died by 1950.

(b) The firms that are born are smaller in size and have a smaller variance than those alive at the end of the period.* The births shown in the period 1924–39 are, exceptionally, of a greater mean size than those alive throughout the period; this is because the births included in this set of columns are only those on the Official list while a large number of other companies were added to the Supplementary list. If the average is calculated for the Official and Supplementary lists together (assuming that the greater part of the Supplementary list in 1939 is composed of births during the period 1924–39) it is found that the mean size of births is reduced to 0·29, and thus conforms with the results observed for the other periods. Since births in this study are largely due to old companies receiving a quotation, no general significance can be attached to the observations for this category of company. Note, however, that births have always decreased the variance of the distributions, and hence have decreased business concentration in the sense shown by Lorenz curves; this is in contrast to the *a priori* argument occasionally put forward that the addition of small companies to the population will generally increase concentration. What seems to happen in the present case is that the smaller degree of concentration among the births dilutes the general level of concentration.†

* Births have to be compared with those alive at the *end* of the period, and not at its beginning, and *vice versa* for deaths, as will be apparent from the method of construction of the basic tables.

† There is a slight paradox here which is worth comment. If more companies become quoted as time proceeds, it would seem that the quoted sector must over time include an increasing number of the smaller size ranges. In that case the variance, and hence the concentration, of the quoted sector ought to increase. Presumably, what happens is that, if a company is to become quoted, it must have a larger market valuation than those in the existing smallest category of companies, which tend to include an undue proportion of unsuccessful companies. In fact, they are so very much larger, that they actually decrease the dispersion rather than increase it.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

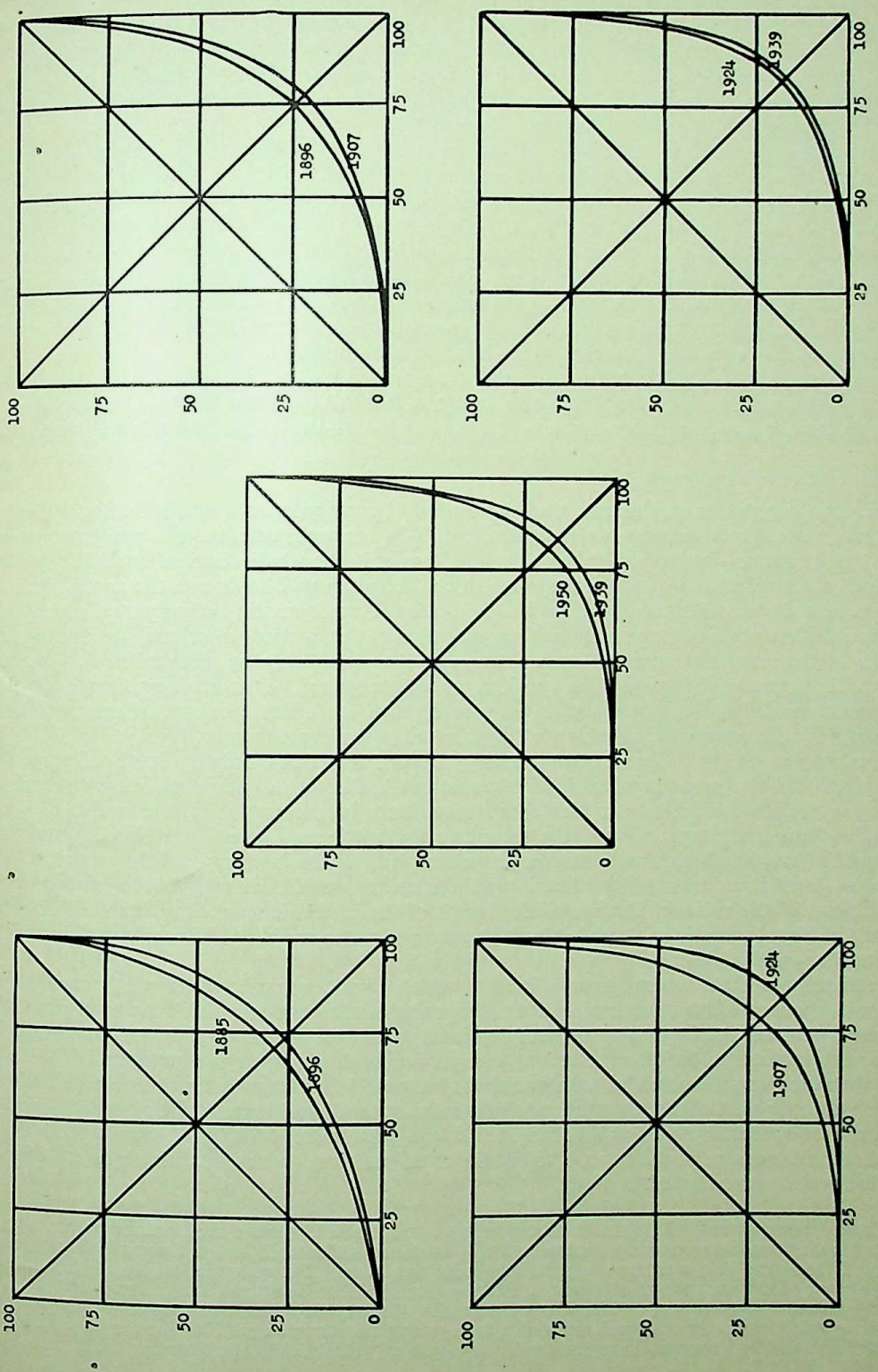


FIG. 5.—Changes in concentration of surviving companies.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(c) The most striking feature of Table 8 is that, despite the large number of births and deaths, a preponderating influence is had by the changes in the sizes of those firms that are alive throughout the period. For all except the last of the periods examined, their variance increases; the amount of change may be assessed with the help of the Lorenz diagrams of Fig. 5, which relate to these constant sets of firms. The decline in concentration in the last period stands in surprising contrast to the earlier periods and receives further analysis below.

(d) A comparison of the magnitude of the short term movements in the Lorenz curves of Fig. 5 with the movements of the calculated variances of Table 8, shows a good agreement. On the other hand, the calculated variances for the total distribution show certain discrepancies when compared with the Lorenz curves of Fig. 1, especially when the extreme dates are compared.† Presumably the reason is that in the short run the main changes are reflected in the variance of the distribution, while the higher moments remain relatively stable; but in the long run the latter may also change substantially as shown by the values of g_1 and g_2 in Table 2 above. It therefore seems useful to supplement the comparison of variances by an inspection of the Lorenz curves before reaching any conclusion on changes in business concentration.

Viewed as a whole, Table 8 shows that there has been a marked tendency for concentration to increase among those firms that remain in existence, but this has been offset by the continuous changes in the population of firms that are alive, as reflected in the birth and death statistics.

4.4. *The Effects of Amalgamations*

When two companies amalgamate the distribution of companies as a whole is affected (a) by the reduction in the total number of companies, and (b) by the replacement of two small companies by one larger company which affects the balance of power in the whole distribution. The greater portion of the amalgamations that take place in the economy will be due to the acquisition by quoted companies of other unquoted companies, or to the amalgamation of unquoted companies (such as may take place prior to their seeking quotation); this portion, of course, lies outside the coverage of the statistics examined here. Amalgamations will only be reflected in the present statistics if both the acquiring and the acquired company are quoted. It is, however, the spectacular amalgamations of large companies that may be expected to affect the level of concentration significantly, and these are presumably well represented in the quoted sector.

In order to assess the influence of such amalgamations a further tabulation has been prepared of all quoted business units involved in amalgamations in the various periods of time covered in this inquiry. For each period two size distributions have been constructed, the first relating to that before their amalgamation, and the second to that after amalgamation. The size measure used for the first distribution was the market valuation as at the beginning of the period. For the second distribution it would have been possible to use the market valuation of the joint concern as at the end of the period; in view of the length of the period it seemed preferable simply to sum the individual valuations at the beginning of the period, so avoiding the problem of deflation for price changes.

The distributions are summarized in Table 9, which gives the number of firms, the mean and variance of the logarithms of their sizes at the beginning of the period in which the amalgamation took place, and the corresponding values after amalgamation. In the period 1885–96 no amalgamations are recorded for this sector. A comparison with the aggregate statistics of Table 8 shows that the firms involved in amalgamations were of larger than average size even before amalgamation; and, further, that the rate of amalgamation hardly changed throughout, about 7 per cent. of quoted firms being involved in a fifteen-year period. The bottom row of Table 9 show that the dispersion of the firms involved in amalgamations was reduced in three out of the four time periods investigated.

† Thus, the variances for 1885 and 1950 are rather close together, whereas the Lorenz curves are well separated. If a sampling error is attached to each of the variance estimates, based on the usual assumptions, it is found that these discrepancies are not statistically significant. For example, taking $\sigma^2\sqrt{2/n}$ as the sampling error of the variance, this gives a standard error of 0·77 to be attached to the variance of 4·17 for 1885, and a standard error of 0·16 for the variance of 4·07 for 1950. The small number of observations in the first period leads to non-significant results in any comparisons made with it; for the later periods, on the other hand, the changes in concentration are generally statistically significant at the 5 per cent. level.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 9

Statistics of Firms Involved in Amalgamation

Units: $\log_2$

	1896–07		1907–24		1924–39		1939–50	
	Before	*After*	*Before*	*After*	*Before*	*After*	*Before*	*After*
Number	11	5	40	13	58	25	109	50
Mean	1·18	2·20	1·15	3·54	1·59	3·24	1·93	4·00
Variance	5·47	6·56	5·18	3·01	6·30	5·06	6·93	4·52

In order to put these figures into perspective it is necessary to superimpose them on the distribution of the remaining companies that were not involved in amalgamations. Using equation (3) it is possible to subtract the distribution of firms that were involved in amalgamations from the total of all firms, and to add back their distribution after amalgamation. In this way an estimated variance of the distribution after amalgamation was obtained and is given in row 2 of Table 10; row 1 of this table is taken directly from Table 8 to provide a basis for comparison. It will be seen that the variance of the distribution, and hence the level of business concentration, certainly tends to increase in consequence of the amalgamations; but the amount of change is quite small in comparison with the changes due to growth given in Table 8.*

TABLE 10

Effects of Amalgamation on the Variance of Firm Sizes

Units: $\log_2$

	1896–07	1907–24	1924–39	1939–50
Total at beginning	3·85	3·79	5·61	6·01
Total after amalgamation	3·85	3·86	5·69	6·13

It appears, therefore, that in studying changes in business concentration over time more attention should be paid to the internal growth of firms than to the possibility of combination among firms.

4.5. *The Growth of Firms*

It is next necessary to examine more closely how the growth of firms affects business concentration over time. The relevant data are those in the central portions of Tables 3 to 7, which are here summarized further in a diagrammatic form.

The first diagram is intended to help in deciding whether it is broadly true that all firms had the same chances of growing or diminishing in size. An inspection of any of the tables, column by column, shows that in each column there is generally a large central element—corresponding to the fact that most firms remain in the same size-class during the period of time considered—and that on each side of that element the frequencies tail off fairly symmetrically; that is, the probability of a given proportionate change in size varies inversely with the magnitude of the change. Fig. 6 shows this rather more clearly. It has been compiled from the combined data relating to the periods 1907–1924 and 1924–1939, the reason for combining the data being that it is necessary to have a large sample of observations in order to diminish the otherwise excessive sampling variability at the tails of the distribution. The small changes in the average size of firms between these periods do not appear to present a problem for the present comparisons.

The firms have been grouped into three approximately numerically equal classes: small (classes A to F), medium (classes G and H), and large (classes J to S). By summing diagonally across Tables 5 and 6 the probabilities of growth of each of these categories have been calculated and are shown in the diagram. It will be seen that the resulting curves are fairly close to one another. But if the horizontal scale had shown the amount of growth in *absolute* terms rather than as a proportion of the original size, it is apparent that the curves would be rather far apart, since the

*These results conform with the more detailed examination of mergers in the United States by Lintner and Butters (1950).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

large firms (classes J to S) are of the order of a thousand times as great as the small firms (classes A to F).

This exercise thus illustrates the operation of the *law of proportionate effect* which, as has been pointed out in section 3.4, was suggested by Gibrat as the explanation of the observed distributions of the sizes of firms.

In addition it suggests a number of new hypotheses, for it will be observed that the tables referred to are essentially bivariate frequency distributions, one of the variates being the size of the companies at the beginning of the period. Since the marginal distributions are approximately log-normal, it may be inferred that they are approximately bivariate log-normal distributions. One of the properties of such distributions (when transformed to normality) is that any cross-section taken parallel to the axes will be normal and, further, that the derived distribution

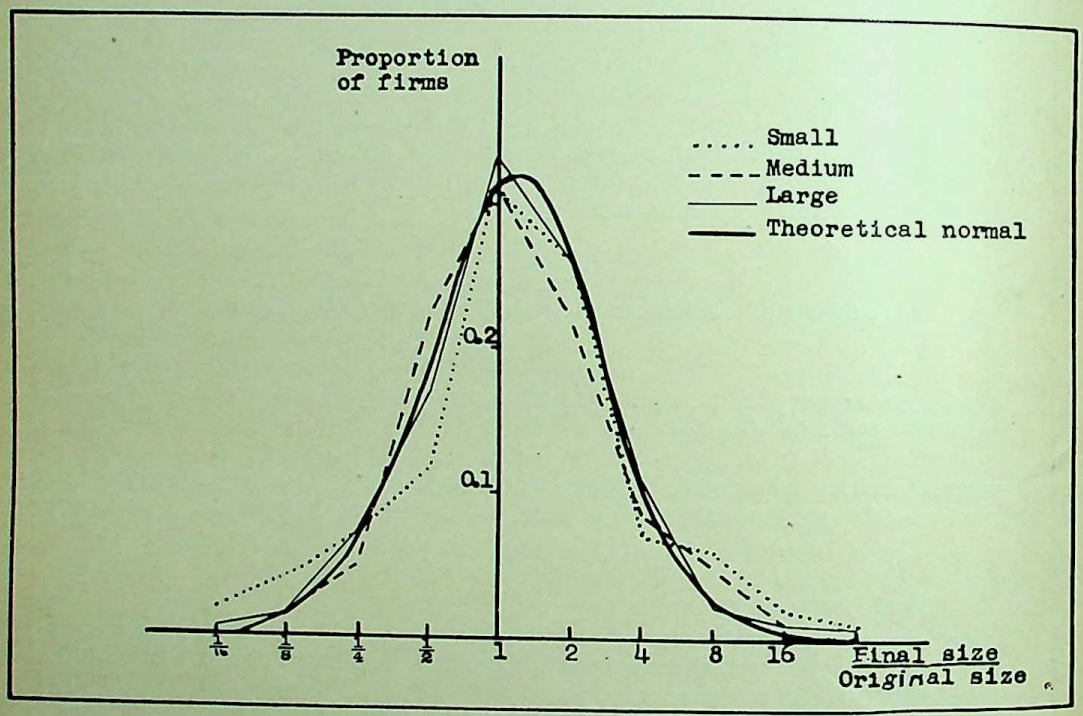


FIG. 6.—Growth of small, medium and large firms during a 16-year period (average of 1907–24 and 1924–39)

obtained by summing diagonally across the distribution will also be normal (Yule and Kendall, 1945 ed., p. 233). Summing parallel to the leading diagonal gives the average probabilities of growth for all firms; it is therefore not altogether surprising that the curves in Fig. 6 have the characteristic normal form. To show the closeness of fit, a normal curve has been drawn in with the same mean, and passing through the maximum frequencies of the combined distributions shown in the diagram.*

The story of the relative fortunes of firms as related by such growth curves is shown for each of the time periods in Fig. 7. The average amount of growth, as shown by the position of the maximum ordinate of each curve, varies from period to period but, as has been pointed out above, this is partly a reflexion of changes in prices and not too much attention should be paid to it. Rather more striking is the similarity of the dispersions of the curves over periods of time which had such different economic characteristics; there are some differences in the lengths of the periods,

* This method of fitting is really no more than a "fitting by eye" and gives most weight to the central frequencies; but it seems adequate for the present purpose. The text-book method of fitting a normal distribution, by equating sample to population moments, gives an unsatisfactory fit in this case to the centre of the distribution, though presumably the fit is better at the tails.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

but these do not seem to be sufficiently great to affect the general result. The slightly larger variance of 1924–39 is noticeable and can, perhaps, be attributed to the great depression. Again, the symmetry and near normality of the curves over such a wide range of values may be noted. Taken together these curves provide strong evidence for saying that the number of companies which quadruple their size is approximately equal to the number that quarter it, that the number that grow sixteenfold is approximately equal to the number that are only a sixteenth of their original size, and so on;* further, the frequency of these proportionate growths is distributed approximately as the normal curve of error.

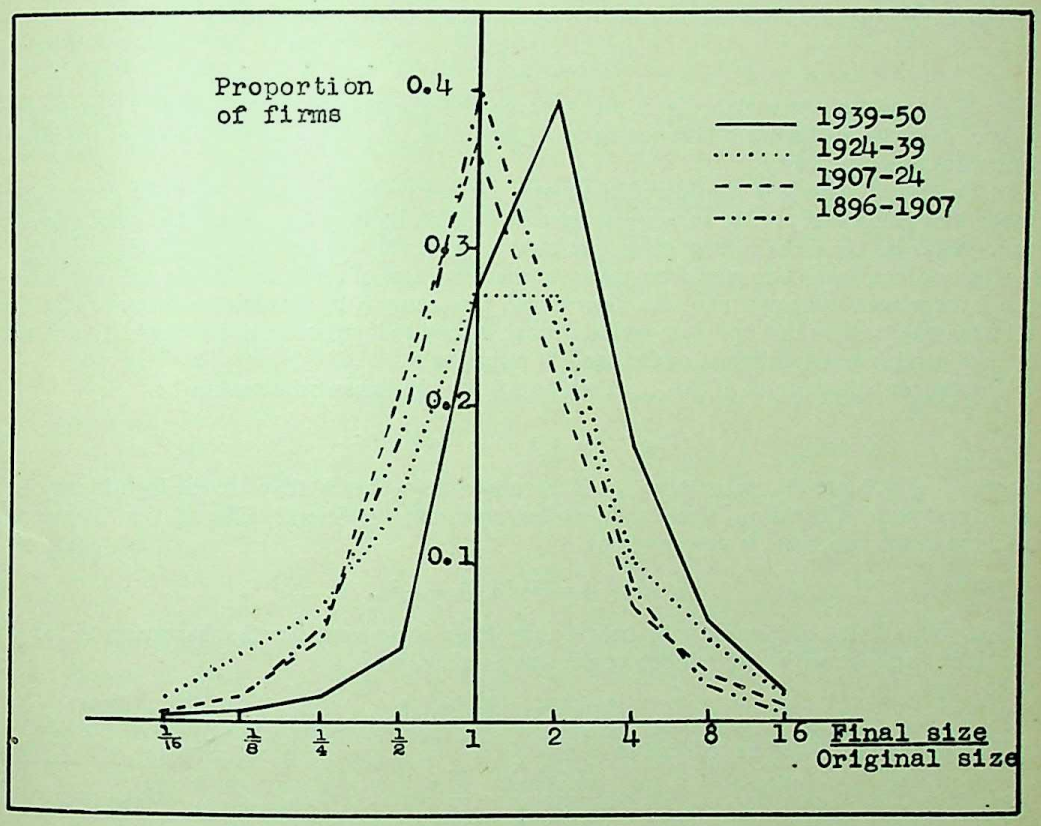


FIG. 7.—Probabilities of the growth of firms, 1896–1950.

4.6. *Growth and Concentration*

If such a law of proportionate growth held without any restriction the consequence would be a continual increase in the dispersion of the sizes, that is to say, the concentration of industry would increase over time. This may be perceived by considering the largest company at any point of time; after an interval there would be a certain probability that its size had increased beyond that limit and, given a large sample of companies, it is to be expected that a certain proportion of them will have grown beyond the previous maximum size. Similarly, at the other end of the distribution, the lower limit to the sizes will over time gradually be reduced. As Galton† remarked in his remarkably analogous discussion of the relation between fathers' and sons' heights (or other hereditary characteristics) "the giants (in any mental or physical particular) would become more gigantic, and the dwarfs more dwarfish, in each successive generation". A possible counteracting

* Examples of companies that show such very large changes in size are given in Appendix D in the form of case studies.

† In his preface to the 1892 edition of *Hereditary Genius* (p. xvii).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tendency which could stabilize the situation would be that of "regression", as it was termed by Galton.*

To explain this, it is necessary to recall the algebra of the theory of correlation. Let x_t and x_{t+1} be the logarithms of a firm's size at a base date t, and any subsequent date $(t + 1)$. Then, without regression, x_{t+1} and x_t would be related by an equation

$$x_{t+1} = x_t + \varepsilon \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where, since we are dealing in logarithms, ε is now an additive random variable normally distributed (and independent of x_t) with, say, a zero mean and a variance of σ_ε^2. If the growth of all firms is governed by (4), the dispersion of firms' sizes at the two dates is given by

$$V(x_{t+1}) = V(x_t) + \sigma_\varepsilon^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where $V(x)$ indicates the variance of x, and necessarily increases over time. After n such intervals of time, in each of which a relation such as (4) held, the increase in the variance of the sizes of firms would then be $n\sigma_\varepsilon^2$.

In the simple case just considered there was no regression present. Suppose next that the change in a firm's size tended to vary systematically with its deviation from the mean size of all firms. That is, the firms in any given size-class at t would still be distributed log-normally at $(t + 1)$, but their mean size would be nearer to the mean size of all firms. Such a situation could arise if there were some optimum size determined, for example, by financial criteria, so that firms that were smaller than this optimum would over time tend to increase in size towards it; and firms that had exceeded this optimum would tend to return to it.

The simplest illustration of this is to replace (4) by the linear relationship

$$x_{t+1} - \bar{x}_{t+1} = \beta(x_t - \bar{x}_t) + \varepsilon \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $\bar{x}_{t+1}$ and $\bar{x}_t$ are the means of x_{t+1} and x_t respectively (that is, they are the logarithms of the geometric means of the firms' sizes) and β is the regression coefficient. The relation between the variances at the two dates is now given by

$$V(x_{t+1}) = \beta^2 V(x_t) + \sigma_\varepsilon^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

so that with a value of β sufficiently below unity there is no need for the dispersion to increase over time. In fact, with a correlation coefficient ρ defined as

$$\rho^2 = 1 - \sigma_\varepsilon^2 / V(x_{t+1}) \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

it will be seen that (7) can be reduced to

$$\frac{V(x_{t+1})}{V(x_t)} = \frac{\beta^2}{\rho^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and the variance, and therefore the degree of business concentration, will decrease if the regression coefficient is less than the correlation coefficient. If the regression coefficient is greater than unity, however, the variance will always increase since, of course, $\rho \leqslant 1$. It will be noted that the least squares estimates of β and ρ satisfy (9) identically.

The concepts introduced in the previous paragraphs which govern the changes in business concentration can now be related to the general notion of the size mobility of firms. In order to assess the state of competition it is important to have some measure of the rate at which firms change their relative position in the size distribution; a rank correlation coefficient between their positions at two dates would provide one measure, another would be the usual product-moment correlation coefficient. The latter coefficient gives more weight to a large change than to a small change in size, and this presumably is an advantage. There is, however, a conceptual disadvantage in such measures of mobility; for the number of interchanges in the ranks of firms depends on the number of firms and how densely they are packed together. Thus with a given rate of variation in size, a period (or country) in which the firms were originally all widely separated in

* For discussions of Galton's problem in an economic context, cf. Kalecki (1945); the remarks by Marschak (1946), p. xvi; and, more generally, the discussion by Friedman (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

size would show fewer interchanges than another in which the sizes of the firms were closer together.

For this reason it is convenient to use a measure of mobility which is independent of the original distribution of firm sizes; such a measure is provided by the residual variance of the growth equation, denoted by σ_ε^2 in (7). This measure corresponds (after correcting for regression) to the variance of the growth curves of the kind shown in Fig. 7 above. A value of $\sigma_\varepsilon^2 = 1$ would show that a set of firms that were of the same size at the beginning of a period of time would, at its end, have a variety of sizes such that (assuming normality) approximately 16 per cent. would be more than double the mean size, and about 16 per cent. would be less than half the mean size; or, let us say, 32 per cent. had "substantial" changes in size. As the variance increases the proportion having a "substantial" change in size increases; thus for $\sigma_\varepsilon^2 = 1{\cdot}25$, $1{\cdot}5$, $1{\cdot}75$ and $2{\cdot}0$, the corresponding proportions are 37, 41, 45 and 48 per cent. (values for any other degree of change in size may be found from tables of the normal integral).

In order to understand the effects of the growth of firms on the concentration of industry it is thus necessary to know the magnitude of the dispersion of the growth curves, and of the regression coefficient. These have been calculated for the five periods investigated, and are given in the first two rows of Table 11. The third row gives the standard error of the regression coefficient and provides a measure of the limits within which the true value of the regression coefficient may lie; the fourth row gives the correlation coefficient.

TABLE 11

Variance and Regression Parameters of the Concentration Process

		1885–96	1896–07	1907–24	1924–39	1939–50
	Calculated values:					
1	s^2	2·40	1·42	1·82	3·08	1·31
2	b	0·95	0·91	1·09	0·92	0·75
3	$s(b)$	±0·15	±0·05	±0·03	±0·03	±0·02
4	r	0·83	0·89	0·92	0·88	0·85
	Standardized values:					
5	$\bar{s}^2$	2·18	1·29	1·04	2·01	1·24
6	$\bar{b}$	0·95	0·91	1·05	0·95	0·77

Since the time periods vary in length, and since over a longer period of time there is greater scope for changes in the sizes of firms, it is necessary to standardize these estimates on to the basis of a common time period; they may then be compared with one another. In rows 5 and 6 of the table are given estimates of the residual variance and the regression coefficient standardized to a common ten-year period according to a method described in Appendix C.

For the four periods to 1939 it is found that the regression coefficients are in the region of unity; for the two periods before 1907 the sampling error is too large to be confident as to whether they are greater or less than unity, but for the two periods after 1907 it seems that the coefficient was somewhat larger than unity for the first of these, and somewhat smaller than unity for the second. In the final period 1939–50 the regression is very much lower than unity and shows that there has been a clear break with the past. Thus the decrease in concentration observed for the last period is seen to be a consequence of a tendency for firm sizes to regress towards the mean; a tendency which appears to be of substantial importance only in this period.

The growth variances given in row 5 show that the size mobility of firms, and therefore, the intensity of competition, has varied considerably over the time periods considered here. On the average it appears that σ_ε^2 has a value of about $1\frac{1}{2}$, indicating that the typical experience of British industry in the past half century was that, over a ten-year period, about 21 per cent. of firms double their size or do better, about 5 per cent. (which are included in the 21 per cent.) will quadruple their size or do better, and so on; and similar proportions apply for those which halve and quarter their size or worse. During the depression of the '30s, which falls into the fourth period examined here, mobility was rather greater than this; and the same is true for the first period which included the depression of the early '90s.

A final question that requires investigation is the nature of the change in regression, and the reversal of the trend in concentration, in the last period. Could this be due to some imperfection

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in the calculation, or must it be taken as representing a true change in the state of affairs? Fig. 8 is a scatter diagram for this period as calculated from the central portion of Table 7; that is, the average sizes of firms in 1950 are shown for each given 1939-class. These averages have been calculated in two ways. First, the means of the logarithms of the sizes were taken, and these points are shown in circles; these form the basis of the calculated regression in Table 9. Secondly, an arithmetic average was calculated, including the firms that had died by 1950 by giving them a size of zero; these are shown in the figure, when they are graphically distinct, as crosses.

The scatter of points reveals a remarkable curvilinearity, with the larger firms all having much the same proportionate increase in size and the smaller firms showing a considerably greater than

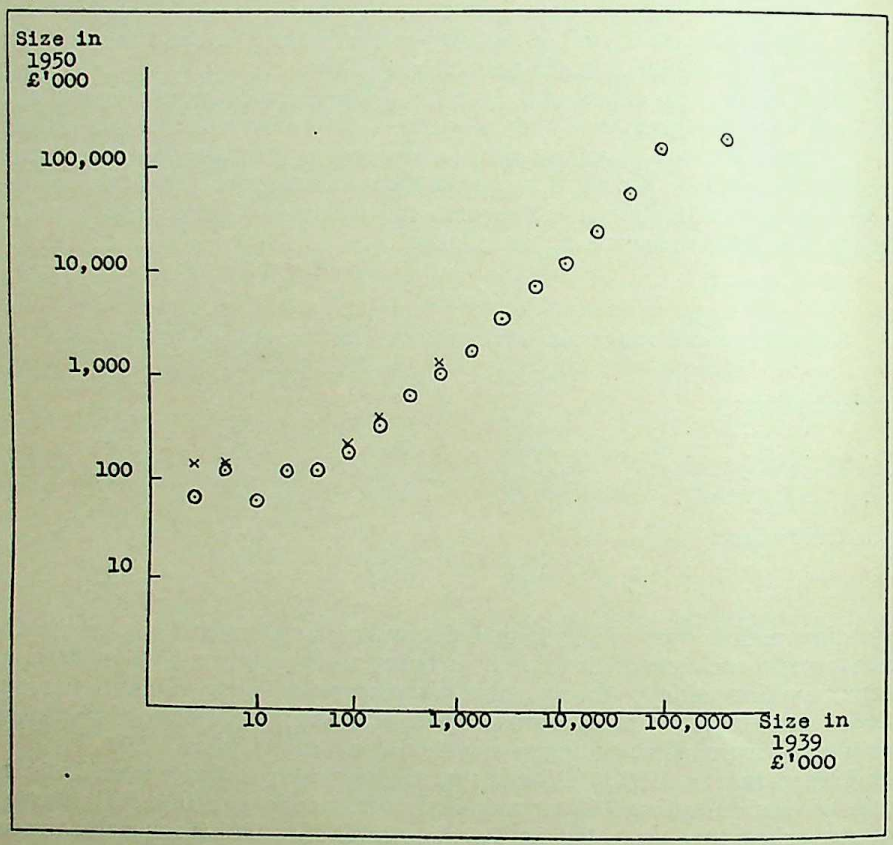


FIG. 8.—Regression of firm sizes, 1939–50 (logarithmic scales).

average increase in size. That is, it is the smaller firms that regress towards the mean. The calculation of the second set of points (the arithmetic averages) shows that this is not merely due to the deaths being ignored in calculating the regressions. The arithmetic mean is higher than the geometric mean (by an amount that depends on the variance of the distribution) if the same set of firms is considered, but is lowered in the present calculation by the inclusion of deaths. It follows that regressions based on arithmetic means would differ from those calculated for Table 11, but this diagram suggests that the differences would not be very great. Similar diagrams have been prepared for the other periods but have not been reproduced here; in these the regressions do not exhibit any noticeable curvature.*

It is tempting to conclude that the profitable experience of the smaller firms in the period 1939–50 was exceptional and, hence, that the overall absence of any substantial degree of regression

* The possibility that the regression in the last period may merely be due to changes in the proportion of capital quoted on the Stock Exchange is examined in Appendix B. It appears that this cannot account for the regression.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

towards the mean is the main conclusion to be drawn from the present analysis. In that case, it would appear that throughout the period investigated the forces determining the growth of firms have been such as to increase business concentration in the manner suggested by the simple process (4). It is only the continual appearance of new firms which mitigate this increase, and this investigation thus points to the importance of maintaining conditions for the free entry of new firms. If for any reason new firms ceased to appear, concentration would increase at the rates suggested by Fig. 3 for the period up to 1939.

5. CONCLUSIONS

The more general aspects of our investigations may be summarized as follows.

1. The central conclusion to be drawn from this inquiry is that the concentration of industry has tended to increase among surviving firms in the U.K. in the last half century simply as a result of the laws which equally govern the growth of small and large firms. If there were some optimum size of firm to which all firms tended, or if it were the case that giant firms owed their continued existence to men of exceptional ability who could not be replaced, then it would be possible for the level of concentration to remain stable. The present investigation shows that no such "regressive" tendencies were present in the U.K., at any rate up to 1939. For the period covering the second World War, the smaller companies seem to have had an exceptionally profitable record; their greater rate of growth has had the effect of lowering concentration between the years 1939 and 1950 that have been examined here.

2. The rise of new industries and new firms may offset this general tendency for concentration to increase; this is shown in the present investigation by the effects of those companies that receive a quotation for the first time on the Stock Exchange. Throughout the periods examined, the addition of these companies leads to a lowering of the degree of concentration in the quoted sector.

3. Consequently, changes in business concentration in the economy as a whole over the past half century may not be very great. Of course, only the quoted sector of the economy has been examined here since our concern has been more with the *forces* determining concentration than with the actual level of concentration; further calculations on the economy as a whole are in progress and will be reported in due course.

4. The principal tool of analysis has been the log-normal distribution which has been investigated by Kapteyn, Gibrat, and Aitchison and Brown. The simple two parameter form has been fitted here to the sizes of firms; while statistical tests show that there are significant deviations from the theoretical curve, it has nevertheless proved to be a useful model as a basis for further investigations.

References

AITCHISON, J., & BROWN, J. A. C. (1954), "On criteria for descriptions of income distributions," *Metroeconomica*, *6*, 88–107.

BERLE, A. A., & MEANS, G. C. (1932), *The Modern Corporation and Private Property*. New York: Macmillan.

BLAIR, J. M. (1950), *Statistical measures of concentration in business: Problems of compiling and interpretation.*" Paper delivered to Amer. Stat. Assoc. convention (mimeographed).

COPEMAN, G. H. (1955), *Leaders of British Industry*. London: Gee.

CROXTON, F. E., & COWDEN, P. J. (Undated), *Applied General Statistics*. London: Pitman.

EDWARDS, C. (1949), *Maintaining Competition*. New York & London: McGraw Hill.

FISHER, R. A. (1941), *Statistical Methods for Research Workers*. 8th ed. Edinburgh: Oliver & Boyd.

FRIEDMAN, M. (1955), Contribution to discussion on pp. 230–8 in *Business Concentration and Price Policy* N.B.E.R. Report, Princeton.

GALBRAITH, J. K. (1952), *American Capitalism*. London: Hamish Hamilton.

GALTON, F. (1925 Reprint), *Hereditary Genius*. London: Macmillan.

GIBRAT, R. (1931), *Les Inegalités Economiques*. Paris: Sirey.

GINSBERG, M. (1929), "Interchange between social classes", *Econ. J.*, **39**, 554–565.

GLASS, D. V. (1954), Ed., *Social Mobility in Britain*. London: Routledge.

KALECKI, M. (1945), "On the Gibrat distribution," *Econometrica*, **13**, 161–170.

KAPLAN, A. D. H. (1954), *Big Enterprise in a Competitive System*. Washington: Brookings Institution.

KENDALL, M. G. (1943), *The Advanced Theory of Statistics*. London: Griffin.

LEAK, H. & MAIZELS, A. (1945), "The structure of British industry," *J.R. Statist. Soc.*, **108**, 142–199.

LINTNER, J., & BUTTERS, J. K. (1950), "Effect of mergers on industrial concentration, 1940–1947," *Review of Economics and Statistics*, **32**, 30–48

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

MARSCHAK, J. (1946), Preface to *Changes in Income Distribution during the Great Depression* by H. Mendershausen. New York: N.B.E.R.
MASON, J. I. & SACHS, D. (1955), "Statistics bearing on financial investment," *J.R. Statist. Soc.* A, **118**, 67–100.
MENDERSHAUSEN, H. (1946), *Changes in Income Distribution during the Great Depression.* New York: N.B.E.R.
NUTTER, G. W. (1951), *The Extent of Enterprise Monopoly in the United States*, 1899–1939. Chicago: University of Chicago Press.
PRAIS, S. J. (1955), "Measuring social mobility," *J.R. Statist. Soc. A.*, **118**, 56–66.
PREST, A. R. (1948), "National income of the United Kingdom, 1870–1946," *Econ. J.*, **58**, 31–62
Register of Defunct and Other Companies. London: Skinner.
ROSENBLUTH, G. (1954), "Industrial concentration in Canada and the United States," *Can. J. Econ. & Pol. Sci.*, **20**, 332–346.
—— (1955), "Measures of concentration," in *Business Concentration and Price Policy*, N.B.E.R. Report, Princeton.
RUTHERFORD, R. S. G. (1955), "Income distributions: new model," *Econometrica*, **23**, 277–294.
SMITH, K. C. & HORNE, G. F. (1934), *An Index Number of Securities*, 1867–1917. L.C.E.S. Special Memorandum, No. 37.
Stock Exchange Investment List. Printed for private circulation by F. C. Mathieson, London.
Stock Exchange Official Year Book. London: Skinner.
TRIFFIN, R. (1941), *Monopolistic Competition and General Equilibrium Theory.* Harvard University Press.
TUCK, R. H. (1954), *The Economic Theory of Rank.* Oxford: Blackwell.
WOYTINSKI, W. S. (1943), *Earnings and Social Security in the United States.* Washington: Committee on Social Security, Social Science Research Council.
YULE, G. U., & KENDALL, M. G. (1945), *An Introduction to the Theory of Statistics.* 13th ed. London: Griffin.

APPENDIX A

The Collection of the Basic Statistics

The sector of the economy examined in this paper consists of those companies engaged in mining, manufacturing and distribution which are quoted on the London Stock Exchange. Since we were mainly interested in measuring concentration in the productive sector of the economy, only those companies listed in the categories Breweries and Distilleries, Commercial and Industrial, and Iron, Coal and Steel have been considered, though companies in other categories could be subjected to a similar statistical analysis. Companies operating mainly overseas, whether British owned or not, were excluded throughout on the grounds that the aim was to measure concentration in this country.

The companies which operated mainly in the United Kingdom and which were quoted in these categories were grouped into business units, as defined in section 3, and their sizes were measured by the market valuation of their quoted capital. Capital in this context includes debentures, and these, together with ordinary and preference capital, were evaluated using the selling prices quoted at the date nearest the end of March for the chosen years, in conformity with the practice of the Stock Exchange Records Office. The years 1907, 1924 and 1950 were chosen to enable comparisons to be made with movements of concentration which may be calculated from Census of Production material. The reason for choosing 1939 was that the Stock Exchange Records Office made its first calculation of the total market valuation of securities quoted in the Supplementary List in that year. The years 1885 and 1896 were also investigated to see whether the oft-cited amalgamation movements of the last decade of the nineteenth century occurred within the quoted sector and increased business concentration there.

Nationalisation provided a problem for the last period, 1939–50. It would have been possible to include those companies that were affected and to show them either as deaths or at their new market valuation in 1950: even though at that date they may only have been holding government stock or entering entirely new realms of business. Both of these treatments have their difficulties, and it seemed simpler and more satisfactory to exclude nationalized companies. Consequently,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the Iron, Coal and Steel category was omitted in this period; that is, calculations given in this paper for the period 1939–50 relate only to the Commercial and Industrial, and the Brewery and Distillery sections. It is perhaps worth adding that a duplicate set of calculations was carried through treating nationalized companies as deaths, with virtually unchanged results in all the derived statistics.

The size distributions for the selected years in the period 1885 to 1939 were compiled by a complete enumeration of business units given in the Stock Exchange Investment Lists printed by Mathieson's. These are more convenient to use than the lists issued by the Stock Exchange, the only significant difference being that the former exclude some companies which technically have an official quotation even though there are no dealings in their shares. The size distributions for 1939 and for 1950 as given in Table 1, together with the transition matrix of Table 7, were compiled partly on a sample basis in view of the very large number of companies involved.

The method of constructing the distributions for Table 7 was as follows. The distributions of (a) the largest 200 business units in 1939, and of (b) all other business units in the Breweries and Distilleries category, and of (c) all births and deaths, were obtained by complete enumeration; but the distribution of the remaining 1,220 companies in the Commercial and Industrial category was estimated by taking a one in three sample. In preparing the table, the sample frequencies in each size-class were multiplied by three, added to the frequencies in the other Stock Exchange categories, and the result combined with the distribution of the largest 200 business units.

The slight variations between the closing distributions of one period and opening distribution of another are due to the exclusion of companies that have been reclassified in the second period to a Stock Exchange category outside those considered here. This happens, for example, if in the period 1907–24 a company is concerned with the construction of buildings and at some date in the period 1924–39 changes its activities to the collection of rents from those properties.

APPENDIX B

Some Tests of the Reliability of the Data

1. To the extent that the prices of a parent company's shares reflect its holdings in its subsidiaries, there is an element of double counting involved in adding all the market valuation of those subsidiaries to the market valuation of the parent. Such double counting may lead to an upward bias in the size distribution of business units, and it was decided to investigate its extent by a sample inquiry for 1950. A sample of fifty business units was taken and their sizes were corrected by subtracting that part of the market valuation of subsidiaries which they owned. The minority part of subsidiary companies' market valuation, which parent companies controlled but did not own, remained as part of the market valuation of the business units. The sample chosen is described more fully in paragraph 3 below. Of the 50 companies in the sample, 11 had 15 quoted subsidiaries, but the parents' holdings in 8 of these 15 subsidiaries were not quoted so that their market valuation had already been excluded. In addition, there were two more subsidiary companies which were technically entitled to official quotations, but these had already been excluded because there were no dealings in their shares.

When the sizes of the remaining 4 business units in the sample had been corrected by subtracting the market valuation of their subsidiary holdings, it was found that in 3 cases the corrected business units remained in the same size-class, and in the one remaining case the business unit was demoted by one class. The total effects of the double counting are shown in Table 12. It will be seen that the aggregate market valuation of these 4 business units is overestimated by some 4 per cent., but that in the total sample the bias is less than 1·5 per cent.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 12

Market Valuation of a Representative Sample of Companies in 1950

£'000

	Original Valuation	*Corrected Valuation*
4 business units with actively quoted subsidiaries	225,250	216,833
46 remaining business units	367,082	367,082
Total	592,332	583,915

It appears, therefore, that the greater part of parent holdings in subsidiary companies tend to be unquoted, and we conclude that the element of double-counting does not result in any substantial bias in the size distribution.

2. For the period 1939–50 it was decided to make a more detailed study of the smaller companies which showed such an extraordinarily rapid rate of growth in this period (see p. 174 above). In using market valuation as a measure of size only those categories of capital that are quoted on the Stock Exchange are included in the valuation. Quite frequently, however, there are other categories of shares (such as non-quoted debentures, or workers' shares) which could be added to the market valuation—though at a substantial cost in research time—if some conventional method of valuing these shares is adopted. For the smallest companies it seemed worth investigating the effect of these unquoted issues, though a comprehensive investigation for all companies would have been unduly expensive.

There were 21 companies (20 Commercial and Industrial, and 1 Brewery) in the size-classes A–C in 1939 which survived to 1950; these correspond, after correcting for the sampling fractions, to the 61 companies shown in these classes in Table 7. In valuing their unquoted capital the following rules were adopted: if part of the same issue was quoted, the published prices were used also for the unquoted portion; otherwise, debentures were valued at par, preference shares were capitalized at a dividend yield of 6 per cent., and ordinary shares at a yield of 8 per cent. for both 1939 and 1950. The results are shown in Table 13 below. It will be seen that in terms of market valuation these small companies have a relatively large proportion of unquoted capital, in fact, largely composed of unquoted debentures. However, their growth in size between these dates cannot be explained by any increase in the proportion of nominal capital which became quoted, for this proportion decreased slightly over the period. The remarkable fact shown in this table is the growth of the market valuation of the quoted capital by more than tenfold. This growth appears to be due to the rise in the profitability of these firms: information available for 15 of these companies on Moody's cards show that profits earned for distribution increased from an aggregate loss of £56,000 in 1938–9 to a profit, after tax, of £216,000 in 1949–50.

TABLE 13

Estimated Valuation of Quoted and Unquoted Capital of a Sample of 21 *Small Companies,* 1939 *and* 1950

£'000

	1939	1950
1. Nominal value of quoted capital	2,252	2,157
2. Nominal value of unquoted capital	537	715
3. Total nominal	2,789	2,872
4. Market value of quoted capital	187	2,105
5. Estimated value of unquoted capital	375	616
6. Total market valuation	562	2,721

3. For the last year examined in this study, there was available at the Institute the series of standardized accounts prepared in connexion with its inquiry into the finance of public companies. It was thus possible to compare the market valuation of a firm with a number of other measures of size that can be derived from these accounts and to obtain an impression whether the results

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of this paper are likely to be sensitive to the measure chosen. The comparison was based on a representative sample of 50 companies which were selected at random from the 150 companies chosen by Moody's as a basis for their index of company earnings. The following are the comparisons made, together with the correlation coefficients that were obtained:

Market valuation in 1950 with:	R
1. Total fixed assets, net of depreciation	0·80
2. Nominal value of total share capital and debentures	0·87
3. Net assets (i.e., 2 + reserves)	0·86
4. Profits for 1950, net of depreciation	0·89
5. Average profits for the five years 1950–4	0·94

It is rather striking that the correlation with the average profits for a succeeding period of years should be higher than that with the current year. It is difficult to say whether this is because profits for a single year provide a rather volatile measure of size, or because market valuation is based on a realistic appraisal of future trends.

Whilst it is clear that the market valuation of a firm must be influenced by a great many factors (for example, by its distribution policy), it would be surprising in view of these high correlations if different results were obtained using these alternative measures of size. Two measures of size are conveniently available throughout the whole period 1885–1950, namely, Stock Exchange market valuation and nominal capital. The former measure was adopted because it provides a closer approximation to the current value of a company than does its nominal capital, for the latter is only infrequently adjusted to bring it into line with changed circumstances. Market valuation is not a perfect measure of size: but the size dispersions are so great that they can be investigated in spite of the imperfections of our yardstick.

APPENDIX C

Standardized Rates of Regression for a Constant Time Interval

The method of standardizing the regression coefficients of the growth equation, mentioned in section 4.6 above, is as follows. Suppose that the equation for a unit interval of time is

$$x_1 = \beta x_0 + \varepsilon_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where x_0 and x_1 here represent the logarithm of a firm's size in periods 0 and 1 measured from the mean population value of each period, and ε_1 is the random additive factor for the first period. If a similar relation holds for each period of time it is found that

$$\begin{aligned} x_2 &= \beta x_1 + \varepsilon_2 \\ &= \beta^2 x_0 + \beta\varepsilon_1 + \varepsilon_2; \end{aligned}$$

and, in general,

$$x_n = \beta^n x_0 + \beta^{n-1}\varepsilon_1 + \beta^{n-2}\varepsilon_2 + \ . \ . \ . \ + \varepsilon_n \quad . \quad . \quad . \quad . \quad (11)$$

If a relationship is estimated between the sizes of firms over a period of n time intervals such that

$$x_n = \beta_n x_0 + \eta_n \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

then it is apparent that the relations between these parameters and those of the elementary equation (10) are given by

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$\beta_n = \beta^n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and

$$\eta_n = \sum_{i=1}^{n} \beta^{n-i} \varepsilon_i \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Assuming that the εs are independent in time, have zero mean and a constant variance σ^2, it is found that

$$\text{Var}\,(\eta_n) = \frac{1 - \beta^{2n}}{1 - \beta^2}\,\sigma^2 = \frac{1 - \beta^2{}_n}{1 - \beta^2}\,\sigma^2 \quad . \quad . \quad . \quad . \quad . \quad (15)$$

Equations (13) and (15) can thus be used to standardize any calculated values of β_n and Var (η_n) to a standard period of time.

APPENDIX D

Some Companies with Substantial Changes in Market Valuation

According to the law of log-normal growth as treated in this paper, a small proportion of firms will show extremely large rates of growth and decline. In this Appendix a brief description is given of the activities of a number of such companies in order to provide realistic illustrations to what may otherwise seem an unduly rareified statistical exercise: the companies have been chosen by chance simply by inspecting the transition matrices and finding companies that correspond to the elements furthest away from the leading diagonal.

British Vacuum Cleaner and Engineering Co.—In the period 1924–39 the market valuation of this company increased from £3,000 to £375,000 and this is shown in Table 6 by the element HA. Over the same period the net tangible assets increased from £16,000 to £512,000. This company was formed in 1903 to exploit H. C. Booth's invention of a vacuum cleaner; it was soon involved in litigation against other producers and the heavy costs incurred culminated in a 70 per cent. reduction of capital in 1916.

Difficulties persisted after the 1914–18 war and no dividend was paid between 1919 and 1930. The explanation given in the speeches of Mr. Booth, the Chairman of the company, was the resistance of housewives to new methods of cleaning, coupled with competition from imported cleaners. By the end of the 1920s the prejudices against vacuum cleaners had been overcome and this, together with the increasing conversion of houses to electricity, led to an increased turnover and to a profit of £11,000 by 1929–30. In 1931 an import duty of 50 per cent. was placed on imported vacuum cleaners, and turnover increased rapidly. New premises were acquired and a subsidiary was set up in Canada. In 1939 an extra £200,000 ordinary capital was issued and in this year profits (£59,000 earned for dividends) reached their peak.

Courtaulds.—The period 1907–24 saw the rise of this great rayon making concern. Its market valuation increased in these years from £0·4 million (there were in addition £0·2 million of unquoted debentures) to some £36 million; this company is shown in Table 5 by the element QH. This is an astonishing rate of growth but is fully paralleled by the rise in its output of rayon. Production of rayon had begun in 1906, and in the years 1907–9 an average of 150,000 lb. was produced a year; by 1923 it had risen to 14,050,000 lb. With the subsequent rise in competition, growth after 1924 was at a smaller rate so that by 1939 its market valuation had risen only a further 22 per cent. to £44 million.

Major and Co.—This company showed a decrease in market valuation from some £200,000 in 1924 to the trivial sum of £2,000 in 1939; it corresponds to one of the companies in the element AG of Table 6. The company has been concerned with the distribution and refining of petrol,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and the manufacture of its chemical by-products including dyes and coal tar. It suffered considerably in the depression of the '30s as a result of a price war in petrol distribution; for several years it was making losses in this branch of its activities which appear to have substantially eroded its assets. In order to supplement its profits in other ways, a subsidiary was established in 1931 to assemble and sell wireless sets; but with delays in deliveries the first season's sales were missed, and losses were sufficiently great subsequently to lead to the liquidation of the subsidiary in 1933. By 1936 its dye works had been sold to I.C.I. Throughout the '30s the company had negligible profits; trading profits of £14,000 in 1925 had been reduced to £1,000 in 1938, after a loss of £7,000 in 1935.

With the outbreak of war the company's fortunes changed: a petroleum pool was formed as a result of which the company's income from petrol distribution recovered. Its other activities are also now on a much firmer basis, and in the last five years its earnings for dividends, after tax, have been in the region of £30,000 per annum.

Red Tower Lager Brewery.—This company increased in market valuation from £9,000 in 1939 to £120,000 in 1950; it is shown in Table 7 as one of the companies in the element FC. The company was established in 1933 to manufacture lager beer in this country using a Continental process. Losses were regular up to the commencement of the war; even the imposition in 1936 of a customs duty of £1 per barrel on imported lager did not seem to improve profitability. But the war did: in 1942 earnings for dividends, after tax, had risen to a peak of £34,000, by 1948 they were half of that figure, and by 1950 earnings were only at £5,000. The company has thus had a fluctuating history, the relatively large rise in the market valuation over the period 1939–50 corresponding to a change in the market's assessment of the profitability of the concern.

Trawlers Grimsby.—The market valuation of this company increased from £4,000 to £534,000 over the period 1939 to 1950 and is shown by the element KB of Table 7. Its net tangible assets increased from £102,000 to £1,019,000 over the same period. Working profits decreased steadily from £16,400 in 1934 to a loss of £6,800 in 1939 there being a loss in each year from 1936 to 1939. This trawler company shared the general economic depression of the fishing industry which may be attributed to low prices of fish and to increasing costs as the result of the depletion of home fishing grounds.

The advent of war and increased fish prices resulted in an increase in earnings for dividend distribution from a loss of £8,000 in 1939 to a profit of £6,000 in 1941; but this level was not maintained, average earnings 1942–5 being at the rate of £5,000 per annum. It acquired the capital of Grimsby Motor Trawlers Ltd. and of Parkholme Trawlers Ltd. in 1946: in 1948 it acquired East Riding Trawlers Ltd. and Ross Group Ltd., and issued £250,000 loan stock to the latter's former shareholders. The rapid growth of this company in the period 1939–50 can therefore be attributed to two main factors: the amelioration of depressed pre-war fishing industry and the series of amalgamations in the post-war years.

Discussion on Paper by Mr. Hart and Dr. Prais

Professor Champernowne (in proposing the vote of thanks): Dr. Prais and Mr. Hart have suggested the use of the variance of the distribution of the logarithm of size as a measure of business concentration. This measure has several advantages—some dependent on the claim that the distribution of size is almost exactly log-normal, but others independent of that claim.

The main advantage of this measure, *whatever* the shape of the distribution, is that the causes of its change over time can be analysed into those due to the birth of firms, the death of firms and the change of size of firms in the manner that the authors have shown in Table 8.

The main further advantage that it would have if the shape of the distribution could be relied upon to be log-normal is that it would give us complete information about the Lorenz curve and, given the number of firms, would enable us to derive as a function of it any reasonable measure of concentration, such as the proportion in the total industry represented by the five largest firms.

If, however, we could not be sure that the curve would always be log-normal it does not seem to me to be a very suitable measure of concentration—for in measuring concentration we are chiefly concerned with the end of the right-hand tail of the Lorenz curve—that is to say with the degree of dominance of the few largest firms: and even with a given number of firms there is no

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

guarantee of a close correspondence between this aspect of concentration and the variance of the logarithm of size—unless we can be assured that the distribution *will* always be log-normal.

It is therefore a matter of some interest and importance to examine the claim that these distributions are likely to be almost exactly log-normal.

There are two legs supporting this claim: (i) theory, (ii) facts.

If there were no births and deaths of firms there would, as has been explained, be a very simple Gibrat-type model accounting for a log-normal distribution of ever-increasing variance as in equation (5)—and a less simple Kalecki variant of this model which would lead to a stable distribution.

But the fact that there are so many births and deaths completely destroys this line of reasoning. One would have to make very complicated assumptions about the size distribution of the dying and new firms in order to retain the log-normal property of the equilibrium distribution. Professor Rutherford attempted a model of this kind for income distribution, and claimed that it produced something very like a log-normal distribution but, so far as I can see, the main variant of Rutherford's model leads to an equilibrium distribution which is not log-normal but a Pareto type distribution. Thus, when many firms are being born and dying, the theoretical leg for the log-normal distribution becomes very shaky.

The other leg, on which the log-normal claim rests, is the facts: this leg is far stronger. It is very easy to plot distributions on logarithmic probability paper, and to be impressed by the straightness of the plot and claim the distribution to be log-normal, when in fact it is very far from being so: this is often done with income distributions. But it will be found that the distribution of company profits as shown every year in the Inland Revenue reports does give a remarkably straight line year after year: and there seem to be grounds for claiming that these data can be quite well fitted by a log-normal distribution.

Incidentally, I had the impression from these data that the degree of inequality had been increasing again in recent years. Perhaps the authors can throw more light on this.

The figures in the paper are closely related to profits and also show a fairly good fit to the log-normal curves: whether so good a fit I doubt. Incidentally, is it wise in Appendix B to use the *correlation coefficient* as a measure of the relation between variates whose distributions are so very highly skew? This measure must be dominated by the effect of the few giant firms.

At one point the authors appear to claim that because the Lorenz curve is found to be symmetrical, and because the Lorenz curve of the log-normal distribution is symmetrical, and because they know of no other distributions which have symmetrical Lorenz curves, ergo these distributions by size of firm are log-normal.

I understand that Professor Kendall will discuss this point, but I should like to make one observation now: if $f(x)\,dx$ denotes the number of firms of sizes x to $x + dx$, and we choose *any* form for $f(x)$, $(x \geqslant c)$ we can obtain a symmetrical Lorenz curve by putting

$$f(x) = (c/x)^3 f(c^2/x)$$

$$\text{for } 0 < x < c$$

and c will be the mean size. This condition will be satisfied if

$$f(x) = (c/x)^{3/2}\,\varphi(x + c^2/x) \geqslant 0 \text{ for any function } \varphi. \qquad (0 < x < \infty.)$$

A particular family of curves is when the distribution function for size of firm is

$$F(x) = \{1 + (x/c)^{-\frac{1}{1-b}}\}^{-b} \text{ and } f(x) = \frac{dF}{dx}, \qquad 0 < b < 1$$

when $b = \frac{1}{2}$ the Lorenz curve is a quarter of a circle.

I have the impression that the 1907 data can be moderately well fitted by a curve of this family with $b = \frac{1}{3}$, but I have not checked whether the fit is as good as with a log-normal distribution.

In any case the symmetry of the Lorenz curve is not alone a good reason for supposing the distribution to be log-normal.

Finally, I should like to congratulate the authors of this paper on having combined a theoretical discussion of the size distribution of firms, not only with a study of the actual distributions, but also with a study of the matrices of Tables 3 to 7, showing proportions of firms in various size groups which have moved into the various other size groups. Moreover, on the face of it, these matrices seem roughly to confirm the assumptions required for their model, although the difficulty about births of firms and deaths of firms remains. These tables throw a great deal of light on the process of changes of concentration of business in this country.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Mr. GORMAN (in seconding the vote of thanks): I have found this paper most useful. Professor Champernowne has already made many of the points I had hoped to make, but possibly I can fill in a few gaps. I think the first thing in attempting to measure the degree of concentration in industries is to recognize clearly that no particular measure can be of any particular value in itself. The degree of concentration in a group of industries will depend upon the degree of homogeneity. If the top ten firms in a group of ten industries each control 15 per cent. of the total production in one of the industries the situation is quite different from that in which each controls 1·5 per cent. of the total production in each of the ten industries.

That is the point which Professor Champernowne has made. In any measure of concentration the important thing is that it should give some picture of the type of distribution. In looking at any suggested measure the first thing to do is to say how close the measure approximates to theory, how closely it ties up with measures based on the type of the distribution.

I had some computations done on Table 1 to find how well the log-normal distribution fitted the top tail of the Lorenz distribution. The precise results depended critically on the precise position which the largest firms were assumed to take in their size-groups. However, it appeared that there are about twice as many large firms as would be expected on the log-normal hypothesis. This hypothesis clearly fails to give a good picture of the important upper tail of the Lorenz. curve.

The position is probably not as bad as one might think. In fact we have fitted the log-normal figures to the figures obtained from the Stock Exchange and there are too few firms at the bottom because there is an unnatural cut-off. If you remember that the theoretical Lorenz curve is symmetrical you will easily see why this leads to the number of large firms also being underestimated. If then we are to use the standard deviation of logs as the measure of concentration, some estimate based upon the top of the distribution alone should be used.

I myself found that the actual symmetry of the Lorenz curve was not as great as all that. This can easily be seen from Fig. 1. Some curves which are far apart from each other on one side of the diameter are close together on the other.

There seem to be two difficulties about the theoretical basis for log-normal distribution, one, pointed out by Professor Champernowne, that births and deaths form such a large proportion of the data and the variance due to a change over a period of 10 to 15 years is relatively so small in comparison with the variance of the original distribution itself. The variance of the controls is comparatively large and there is a change of variance due to growth. This suggests that there is no time for the random effect to have much effect on the distribution at the end of the period and it is only because the original births follow a distribution reasonably close to log-normal that one gets something like log-normal distribution at the end.

My last point is the question of amalgamations. One of the snags about this is, that one can quite easily see that if fairly large firms are amalgamated, the effect is small in the first and second moments of the distribution but much greater in the higher moments. It is intuitively acceptable that this should increase the skewness; it also increases the kurtosis. The skewness does appear to have been growing throughout the period.

The choice of the particular measure, that is to say the Stock Exchange valuation used by Dr. Prais and Mr. Hart, is largely based on the fact that it is not a good measure, but it is what we have got and it is fairly well related to what we have not got. Its correlation with other more acceptable measures is about 0·8. The correlation of 0·8 obviously corresponds to a squared correlation of 0·64, so that the measure used may explain less than two-thirds of the variation in a "true" measure. Moreover, it is possible that there is a third variable, uncorrelated with the "true" measure, which nevertheless has a correlation coefficient as high as 0·6 with the Stock Exchange valuation. An analysis based on the Stock Exchange values may therefore be misleading.

The vote of thanks was put to the meeting and carried unanimously.

Mr. COLIN CLARK: I speak as a consumer rather than as a producer of works of this nature. But it seems to me that the authors have discovered a very valuable new technique, which should be capable of application to several other distribution problems. Professor Champernowne has suggested that it should be applied to one problem which interests all of us—the distribution of income. To apply this technique, however, in its full beauty we shall need data not yet available, showing for each individual not only the income-class into which he falls this year, but also the income-class into which the same individual fell in years recently past. A beginning has been made in the State of Wisconsin in assembling such information, and it may return fascinating material for mathematicians and economists. I can see this technique being applied also to the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

gradual but significant changes which are taking place in the size distribution of towns and cities, and also to the rather different problem of measuring consumption as a function of income. The authors have something in common with the work of their colleagues Aitchison and Brown whose book on the use of log-normal distributions, I understand, is just due to be published. The Aitchison-Brown results, which in the opinion of Professor Stone represent an extremely important step forward (they were not published by this Society, but many members may be familiar with them), measured consumption as a function of the log of income. Consumption of any one commodity has a saturation limit, whereas incomes may have to be treated as if they had no natural limit. Between zero and saturation this relationship between consumption and log of income must rise, they postulate, on a nicely-shaped S-curve. This rule, they hold, applies to all commodities, though of course we are at very different positions on the curve in respect of bread and in respect of washing machines.

There are a number of different mathematical functions which can generate such a curve, as well as the probit which they favour. I would ask that an ingenious mathematician, interested in these problems, should catalogue possible functions, I hope as simply as possible, to see which best fulfil the conditions, for there is certainly a great deal of practical work waiting to be done in this field. Messrs. Aitchison and Brown's own work dealt only with a single set of data, the Ministry of Labour Inquiry of 1937. As more information comes forward, we see more clearly how many disturbing factors there are. Even for comparatively simple applications, we need a function which takes account of the fact that for certain so-called "inferior goods" consumption actually declines beyond a certain level of income. Consumption also is more affected by prices than is sometimes admitted. In the case of durable goods also it is affected by the stocks of goods already held by the consumer. I have, however, found quite nicely shaped S-curves for the consumption of eggs and of fuel.

Going back to the subject matter of the paper, I think the most important conclusion was that amalgamations were of minor importance in the general picture. This is a highly important conclusion. We have all grasped, I hope, the point that quick rates of growth will, of themselves, lead to an increase in variance. (This rule would also apply in problems of town size and income distribution.) In size of business therefore we must expect increasing variance, unless this factor is offset by an increasing number of births of small businesses. One of the most serious effects to be expected from the present-day rates of taxation is that it may slow down the naturally expected number of births of small new businesses.

The authors claim that their results are not biased by including only shares quoted on the Stock Exchange. But I do not think that they manage to prove it. As some readers may have adverted, there is a serious conflict between the impression which they give, of decreasing concentration between 1939 and 1950, and the results published by the Board of Trade. A distribution table of manufactures by size (measured by number of employees) was published for 1953. Comparing it with 1930 we find a great increase in the relative importance of the larger manufacturing firms. Firms employing over 1,000 are now employing 31 per cent. as against 24 per cent. of the total numbers. Even after discounting the effect of the all-round growth of manufacture since that date, this still represents an increased concentration.

The subject matter of this paper, of course, is not only manufacture. It deals also with commerce, shipping and other firms, although manufacture should have provided the main and indeed the increasing part of the data. It seems to me that there are some unresolved discrepancies in the basic material. I should add the further point that anybody who knows the conditions of present-day business knows that all the best brains in the country are engaged in *not* earning income, but in diverting all possible assets into businesses which can be sold so as to give capital gains in an untaxable form. This may have had some odd effects on Stock Exchange values and their distribution.

If I may be critical, I do see in this paper something which reminds me of the old agricultural proverb, of trying to make a silk purse from a sow's ear. It seems to me the application of an excellent technique to defective material, which needed much more careful consideration. There is often a temptation to do analytical work just because material appears to be available. It is a mistake to think that the materials exist for the purpose of trying out techniques, when the real truth is that statistical techniques exist for the purpose of improving materials. Some people may say that preliminary work on sows' ears constitutes a useful training for students. But even in this case it is better to buy a smaller quantity of more genuine material.

Professor M. G. KENDALL: I join with previous speakers in congratulating the authors on a very substantial piece of research painstakingly prepared and lucidly presented.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

At the beginning of section 3.3 the authors state that the log-normal is the only distribution they know which yields a symmetrical Lorenz curve. It may be useful to give the general family of distributions with this property. If $g(y)$ is an even function of y, then any continuous frequency function $f(x)$ with mean μ and symmetrical Lorenz curve is given (and only given) by the expression

$$f(x) = \left(\frac{\mu}{x}\right)^{3/2} g\left(\log \frac{x}{\mu}\right)$$

provided, of course, that this is an admissible frequency function. This is very similar to Professor Champernowne's formula.

If we put $\log x/\mu = u$, this becomes a frequency function of u, say $h(u)$, equal to

$$h(u) = \mu\, g(u)\, e^{-\frac{1}{2}u},$$

and in the particular case of $g(y) = \exp(-\frac{1}{2}y^2)$, we have, on adjusting the constant,

$$h(u) = \frac{1}{\sqrt{(2\pi)}} \exp\{-\tfrac{1}{2}(u + \tfrac{1}{2})^2\}.$$

This gives us the log-normal as a particular case.

The property of symmetry in the Lorenz curves, then, derives from the special symmetry of the original distribution, and has nothing to do with normality. If $g(y)$ is taken as y^2, u is a gamma variable.

I have another point of some importance. In spite of the fact that the observed distributions are fitted fairly well by a log-normal, I am not convinced that this distribution is the best to provide a theoretical model. It used to puzzle me a lot why the Pareto distribution of incomes was so stable over decades of intense change in social structure, until it dawned on me that perhaps the reason why it was stable was that it *was* in motion; in short, that it was the result of a stochastic birth-and-death process. Professor Champernowne showed us in 1953 (*Econ. Journal*, **63**, 318) how the Pareto distribution could be derived as such a process by fairly simple birth-and-death rules. Other writers have studied the problem and recently Dr. H. A. Simon (*Biometrika*, **42**, 425) has derived a number of other similar distributions from many fields outside economics.

Now the birth-and-death processes may lead to negative binomials or to Fisher's logarithmic series, or, more generally, to the Yule process.

$$f(x) \propto \frac{\Gamma(x)}{\Gamma(x + p + 1)},\ 0 < p < \infty,\ x = 1, 2, \ \ldots$$

They do not, so far as I know, lead to a log-normal. I should therefore like to ask the authors whether they would not agree that the Yule process, or some elaboration of it, may not provide a better model of business concentration and to suggest that it would be profitable to consider how far the process may mimic the log-normal, or vice-versa, in practice.

Mr. Alan Stuart, who regrets not being here tonight, has asked me to make another point on his behalf. If these distributions are, in fact, generated by stochastic processes it is not entirely clear that the usual methods of testing fit between hypothesis and observation are valid. A chi-squared comparison, for example, between the distribution and a fitted log-normal, is based on the assumption that the observed members are distributed multinomially over the cells of the frequency table; and this may not be so for stochastic processes. Likewise the tests of normality applied in Table 2 may be invalidated. The point is an interesting one affecting many kinds of stochastic processes and might repay examination.

Mr. Mason: If the log-normal distribution could be established on the wider field of companies in general, and the quoted section be shown to correlate reasonably within it, on some measure broader than that of the market value of quoted share capital alone, then the aim of providing economic samples for various purposes would come nearer. The capital structure of companies today is very complex. Many private companies are larger than the "tail" of the quoted public ones. Associations up to 49 per cent. of ownership may be greater than is realized. Concentration by absorption of private companies without expanding capital may also have accelerated since the war. The extent of the effect of the operation of the Capital Issues restrictions is unknown. I wonder therefore whether conclusions are yet possible regarding post-war trends even in quoted public companies alone. More analysis of private companies seems to be required especially as regards the curvilinearity shown in Fig. 8.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Can the 1939–1950 variance strictly be compared with the pre-war data (if the latter include the important iron, coal and steel group), without some further correction? Appendix A is a little ambiguous on this point. What would σ^2 be in 1924–39, if the iron, coal and steel section were omitted?

Also, I wonder whether in recent years concentration in the lay sense cannot have occurred, for various political and fiscal reasons, more in the middle groups than the extremes.

Is it not possible that the larger the Gini coefficient recorded, the greater is the likelihood that a larger fraction of the industry is unquoted on the Exchanges? Can this be latent in the textile data for instance? I should expect an abnormal departure from symmetry of the Lorenz curve for this reason. The apparent symmetry concerns the top two-thirds or so of the full Lorenz distributions which might prove to be more "parabolic" in shape, in, say, the bakery trade, as an extreme example. "Output" and "financial" size distributions need not be the same, as the products and profitability may be quite different over the range within the industry.

Turning to the problem of measurement, the inclusion of loan capital is valuable as it reduces the high gearing on the ordinary share capital. Share capital valuations may vary so widely in different directions for different groups of companies in any one year at a date at the end of March. This is in the middle of the principal dividend season when some prices are ex dividend and others anticipatory. It is difficult to accept correlations of capital values with net assets as more than fortuitous. A problem of net asset comparisons for "current" practical sampling work is the fact that they are only available at balance sheet dates, and so cannot strictly be added as at a particular date. As the authors justifiably assert, the spread of public companies is so wide that the objection for large companies is trivial; yet may it not well hold for the smaller and unquoted companies, where valuations of stocks and reserves are extremely variable?

I should like to stress how variable the public company data are, and how difficult is the problem of securing an adequate small sample. I have permission to quote the following facts:

The correlation of the market value of quoted share and loan capital at, say, December 30th with net assets on the latest balance sheet data will always of course be likely to differ considerably from that at March 31st. But a systematic sample of 1 in 25 of the British companies with share capital quoted operating in the home market in the "Commercial, Industrial" section at December 30th, 1955, showed that in 29 per cent. of the companies the market value of *quoted share* capital was less than 40 per cent. of their net asset value from the last balance sheet.

In 21 per cent. it lay between 40 and 60 per cent.; in 26 per cent. between 60 and 100 per cent.; in 18 per cent. between 100 and 150 per cent.; and in 6 per cent. over 150 per cent.

This distribution stresses the variability of the gearing of the loan and other capital unquoted; the anticipatory element in share pricings; and the time lag on balance sheet data; but much more probably the unsafe sample structure, even on 100 companies, to which I should like now to refer.

There were at the time of the sample 2,511 companies with quoted share capital in the "Commercial Industrial" section. The unquoted share capital would add 7½ per cent. to the total quoted, the quoted loan capital a further 13½ per cent., though some of this related to companies without quoted share capital.

Sample	*Average per Company*					*Actual Total*
	A	*B*	*A + B*	*A'*	*C*	
Drawn from . . .	(1–1400)	(1–1400)	(1–1400)	(1–2511)	(1201–1320)	(1–2511)
Fraction . . .	1/25	1/25	2/25	1/25	1	1
Number of companies .	56	56	112	101	120	2511
Preference						
1. Nom. Capital (£000's) .	450	378	414	335	668	349
2. Market Value (£000's) .	450	288	370	332	631	330
2 as % of 1 . .	100	76	89	99	94	95
Ordinary						
3. Nom. Capital (£000's) .	1633	570	1101	1073	2150	796
4. Market Value (£000's) .	4683	1240	2962	2871	5333	2189
4 as % of 3 . .	287	217	269	268	248	275
(4 + 2) as % of (3 + 1) .	247	161	220	227	212	220

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To provide even these figures accurately by small scale sampling was not found possible by random or systematic means.

If a correct measure for stratification could be assured, it would seem that stratification on variable sampling fractions would eventually prove the most fruitful method. The log-normality theory, if substantiated, would thus help greatly. The records of the 2,511 companies were in alphabetical order in blocks of 200, plus one of 111. The preference capital quoted averaged 30 per cent. of total quoted share capital but the range within blocks was from 25 per cent. to 42 per cent. Market values ranged from 196 per cent. to 601 per cent. of nominal capital in the case of Ordinary share capital.

Systematic and block or random samples of 100 gave a fair picture for preference capital, but Ordinary share data thus derived were still too variable for practical use. Samples of 50 or less were highly suspect. Abnormally low market values were given by the block sample C for the middle size groups, so that the distribution by market value of quoted share capital even on 120 companies was by no means apparently log-normal (see data below) though a different scale or larger sample might bring closer relationship. A systematic sample right through the list (*A'*) provided market value, as the percentage of nominal, reasonably well, and the preference data were generally so covered, but the actual Ordinary figures were still 25–30 per cent. away from the actual average for the "Commercial Industrial" group as a whole.

This section of the Official List has, of course, over the years widened its scope as the companies have extended into new products and new industries. This aspect of the spread, especially into the field of private companies also, I think, needs further consideration before any valid conclusions about the economy can be drawn.

On a point made by Mr. Clark, I should like to see studies of the curve $y = e^{-(x+\frac{1}{x})}$ which, though complex, may be a useful form for consumption-income comparisons (x being always $+ve$).

Block Sample C : 120 *Companies*
(Numbers 1201–1320)

Quoted Nominal Capital £ millions	*Number of companies Total*	*Number of Companies whose Quoted Share Capital had Market Value (£ millions)*												
		0–0·11–	·11–	·22–	·45–	·9–	1·8–	3·6–	7·2–	14·4–	28·8–	57·6–	115·2–	230·4–
0–0·049	4	[4]												
0·05–	11	10	[—]		1									
0·1–	16	5	7	[3]	1									
0·2–	40	2	5	14	[13]	6								
0·4–	21		1	3	7	[9]	1							
0·8–	11				1	7	[3]	[6]	2					
1·6–	10						2		[2]	[—]	[—]	[—]		
3·2–	3			1					2				[1]	
6·4–	2													[1]
12·8–	·													
25·6–	·													
51·2–	1													
102·4–	1													
Total No. of companies	120	21	13	21	23	22	6	6	6	·	·	·	1	1

Mr. BLUETT: I should like to draw attention to one or two points on which I have made some notes. I have certain doubts over the basis used here in determining trends in business concentration. If consideration had been given, when formulating the underlying assumptions, to the consequences of changes in the value of money between one period and another it might have shown that the definitions used here might not be acceptable. Again, consideration of the effects of movements in the prevailing interest rate structure might have shown that another approach to the subject on a difference basis would be more suitable. For the past twelve months we have witnessed on the Stock Exchange a rapid change in the structure and level of interest rates. Experience of this period and of similar times shows that the bigger units among those companies

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

whose capital has a Stock Exchange quotation tend to have larger price variations than the smaller units. When "confidence" among investors is high the yield on high grade Ordinary Shares and Stock falls relatively more than the yield on the less marketable (that is smaller issues) of Ordinary Shares and Stock. The converse is also true, that when "confidence" falls the yield on bigger issues is likely to rise proportionately more than that on the smaller issues. Consequently, I have doubts on the validity of using companies with a Stock Exchange quotation as the basis upon which to ascertain the movement in business concentration without making adjustments for these factors.

If it were possible to discover what proportion of the national output was produced by the publicly quoted companies at various times it might be revealed that such companies are now providing a far larger proportion of the national output than they were 50 years ago. My reason for thinking this is the probable trend is that it is the growing companies and the companies in the expanding industries which have to seek Stock Exchange quotations in order to attract the money necessary to finance their growth. The small static concerns, or those industries where the small unit of manufacture functions most satisfactorily, continue as private companies. Again, the incidence of taxation and particularly of death duties is a force increasing the proportion of companies which have a "quotation". This arises from the need for increased funds which families are no longer able to save to provide for the development of their business as they would have done in earlier years. It also arises from the need to sell part of the business in order to pay, or to anticipate the payment of, death duties.

If it was shown that "quoted" companies are producing a constant proportion of the national output and disregarding the effects of the changes in interest rates and of the value of money, would not this analysis of "quoted" companies merely show that as the giants got bigger there was an infusion of smaller companies who received a quotation and so restored the balance of business concentration among the "quoted" companies?

I agree with earlier speakers that to confine the analysis to companies with a Stock Exchange quotation is likely to introduce so many different forms of bias that it may vitiate the conclusion.

The following written contribution was received after the meeting:

Dr. T. Barna: Mr. Hart and Dr. Prais present a mathematical model illustrating the dynamics of business concentration, as defined by them, and then proceed to construct an empirical analogue to that model. The model describes changes between two points of time by means of a bivariate normal distribution of the logarithms of the size of business firms. On the empirical side, the population observed is confined to quoted public companies and size is measured by Stock Exchange valuation.

Apart from the points made by Professors Champernowne and Kendall, who criticized the mathematical form of the model, the significance of the economic analysis embodied in the paper depends on whether (i) the mathematical model is addressed to the relevant economic problem, and (ii) the empirical data are sufficient.

Under (i) the model is designed to describe changes in that part of the population which is common to two periods rather than changes in the whole population. In other words, "births" and "deaths" are exogenous to the model instead of being part of the phenomena explained by it. I think that in assessing the probability of growth as a function of size account should be taken not only of firms surviving a given time interval but also of the probability of "death". Since, as is apparent from Tables 3 to 7, the probability of "death" is negatively correlated with size, the omission of this factor gives the regression coefficients shown in Table 11 a downward bias.

The omission of "death" from the main calculations is probably due to the fact that the array mean of the logarithms of size tends to $-\infty$ when at least one unit approaches zero in size, as the authors assume happens at "death". The alternative calculation which takes account of "death" but is based on the arithmetic mean, and which is indicated by crosses in Fig. 8, may be thought to be inconsistent with the log-normal assumption.

One solution would be to place a small positive numerical value on size after "death". This is not unrealistic since even bankrupt companies have some value. It can be postulated that the firm is indifferent between, on the one hand, remaining constant in size and, on the other hand, having an equal chance of either "death" or of an n-fold increase in size. The discussion of the variance of growth, which occurs in equation (7), suggests that the probability of a four-fold or bigger increase over a ten-year period was small. Hence "death" might be taken as numerically equivalent to a diminution in size to, say, less than one-quarter of the initial size. If some such correction is introduced, it is not certain that Galton's law of regression could still be demonstrated.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Under (ii) observation is confined to quoted companies. These accounted for a smaller proportion of business activity in the nineteenth century than in later years. Further, there is a sharp break when the supplementary list is added to the official Stock-Exchange list. It can be seen from Table 1, that the supplementary list did not include companies over £2 mn. The larger quoted companies are likely to account for all or most business firms of that size, but the smaller ones are likely to account for only a small proportion of businesses of their particular size and there are reasons to maintain that they are not representative of that category. Hence there are reasons to suspect the growth coefficients attached to the smallest classes of quoted companies.

The limitation of Stock Exchange valuation as a sufficient measure of size can be illustrated with reference to the comparison between 1939 and 1950. During that period the gross trading profits of non-nationalized companies increased 3½- to 4-fold. But dividends and interest paid, before tax, increased by only about one-fifth and evidently Stock Exchange values in 1950 were unduly influenced by dividends as against earnings. Yields of equities fell by one-tenth over the period, so that Stock Exchange values in the aggregate increased by only about one-third.

It is thought that the larger companies were proportionately more affected by special taxes on profits and the dividend limitation in force in 1950 than the smaller ones. These factors may account for the apparent fall in the measure of concentration. It is also known that between 1939 and 1950 *The Financial Times* industrial share index, which represents the largest companies, rose less than a broader based index. Since 1950, however, as various restraints on companies' distribution policy disappeared, there has been a compensatory rise in *The Financial Times* index. This, together with the return to more normal competitive conditions between large and small firms, may suggest that the long-term trend towards concentration is continuing and that conclusions should not be based on 1950 stock exchange values.

Mr. P. E. HART and Dr. S. J. PRAIS subsequently replied in writing as follows: May we begin on a note of agreement; we should certainly have been happier if the tables of statistics in our paper related to all firms in the United Kingdom and their sizes were measured by employment or, perhaps, value of sales. But while Stock Exchange material may be an inferior grade of straw, it is the only kind at present available for our purposes and, as it seems to us, can be used to make bricks even if they are a little rough. (In future, similar calculations could be made by the Census of Production authorities.) What is important is to preserve a sense of perspective about their weaknesses and not to rely unduly on any single result. We have little doubt that economists have been too conservative in using Stock Exchange data, though it is often quite sufficient to establish important results: as with all material subject to extraneous sources of variability, a large number of cases must be examined in order to allow the law of averages to work.

The importance of perspective may be illustrated by a point raised by Mr. Bluett and Dr. Barna. They suggest that a divergence in the trends of the prices of market leaders and of average shares may seriously affect our analysis for 1939–50. An examination of the available indexes of industrial ordinary shares shows a rise of 30 per cent. by the *Financial Times* index of 30 leaders, of 28 per cent. by *Moody's* index of 50 companies, and of 29 per cent. by the broader *Actuaries* index; though no very small companies are included in any of these indexes the spread is quite small and suggests that the point is in the nature of a refinement. Our small companies (see, for example, table 13) show a rise in valuation of nearly 500 per cent.; only a very small part of this rise can be due to this divergency in prices.

Similarly, the differential effects of the dividend freeze as between small and large companies must, it seems to us, be small. We have seen no evidence to show that the change in the distribution habits of small and large companies in the period 1939–50 has been substantial enough to have a significant effect on our results. The *a priori* argument advanced above applies only to ordinary share capital, if it applies at all, whereas the market valuation measure includes preference and loan capital. Even if the argument is confined to ordinary share capital, it is by no means certain that investors have been willing to pay much more for the privilege of investing in a company that would avoid conforming to a dividend freeze.

The possibility of analysing our data by means of a Yule distribution, suggested by Professor Kendall, certainly seems interesting. But a satisfactory analysis along these lines may well require better information on births and deaths than is available here. We are also grateful to him and to Professor Champernowne for generalizing the result on the symmetry of the Lorenz curve.

Our treatment of deaths is admittedly not entirely elegant. But if it is granted that our interest should lie in analysing changes in concentration over a decade, then we interpret Table 8 as showing that for certain purposes it is sufficient to examine only the survivors. We then showed (in Table

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

11) that it was the absence of regression that led to the increase in concentration. As Dr. Barna has pointed out, by ignoring deaths in estimating the regression, our values of b have a downward bias as compared with a procedure that gave a positive value to a company's size at death. This only means, however, that we have been conservative in saying that, in contrast to what is observed in other fields, there has been no stabilizing regression here, apart from the period covering the last war. If we had followed the alternative procedure we might well have found some slight evidence of progression, but other experience suggests that the calculated value of b would be unduly sensitive to the precise size assumed for a death; the results would therefore have been uncomfortably arbitrary.

Mr. Clark suggests that certain other statistics show there has been an increase in concentration over the period 1930–53, and he remarks that this is in conflict with our finding for 1939–50. There are reasons for thinking this conflict to be apparent rather than real. First, the post-war figures he quoted are based on the Ministry of Labour's L-Returns and relate to *establishments*, whereas the size distributions published by the Board of Trade for 1930 relate to the number of *returns* (to be distinguished from firms and from establishments) made to the Census of Production office. For 1935 the Census contains size distributions by establishments and by *firms*, defined as establishments trading under the same name. If establishments are considered for 1935, Mr. Clark's concentration ratio based on units employing over 1,000 falls from 24 per cent. to 21 per cent. If firms are taken, the concentration ratio rises to 31 per cent. for 1935, so that the trend may depend crucially on the definition of the unit. In our work we have attempted to measure concentration for *business units*; it is quite possible for concentration in this sense to remain stationary while for technological reasons workers in smaller plants are moved to larger plants.

Secondly, there is an important difference in the years chosen for the comparison. The first portion of his period, that is up to 1939, overlaps with our analysis for 1924–39, and agrees in showing increasing concentration. Further, after this paper was completed, the *Stock Exchange Year Book* for 1955 was published and we were able to extend our analysis to the quinquennium 1950–5. Using similar methods we found that this period showed a considerable increase in concentration; the regression is once again linear, with a value of $b = 0{\cdot}99$ ($\pm\ 0{\cdot}02$), and a growth variance of $s^2 = 1{\cdot}54$ (the full details of the calculations will be reported elsewhere). As suggested in the last paragraph of section 4·6, the war-period thus appears as a distinct exception in the long-term trend towards increasing concentration. Had we chosen to compare concentration for the years taken by Mr. Clark, it may well be that our results could have been in agreement.

Professor Champernowne has reminded us of the limitations of the correlation coefficient in the context of Appendix B. We have therefore recalculated the correlations after taking logarithms and find that the values, taken in the same order as in that Appendix, are: (1) 0·86, (2) 0·92, (3) 0·93, (4) 0·87 and (5) 0·86. These values are not strikingly different from the linear correlations, but it will be seen that the correlation with (5) is now not higher than that with (4).

Mr. Gorman's calculations on the upper tail are presumably in accord with our Table 2. His finding twice as many large firms as expected is arbitrary and depends on where the line is drawn: the higher it is drawn the greater will be the proportionate discrepancy. Nevertheless, we disagree with him in preferring a measure based only on the upper tail. We agree that such a measure has its uses; but had it been adopted in the present study it would have unbearably complicated the regression analysis of section 4. Furthermore, the value of the log-normal hypothesis does not depend entirely on its accuracy in forecasting the sizes of the largest firms, as implied by both Professor Champernowne and Mr. Gorman: even more interesting is the implication of the dynamic analysis that the variance of the size distribution of surviving firms, and therefore any concentration ratio, tends to increase over time, under specified conditions.

Mr. Stuart has raised a fascinating point: presumably it also applies to the classical examples of testing the normality of fathers' and sons' heights.

We omitted iron, coal and steel companies in the period 1939–50 because of nationalization: Mr. Mason asks whether some further correction is not necessary before making a strict comparison with pre-war data. In all strictness, perhaps no comparison is possible; but if one is to be made we do not see any alternative to our procedure. On his further question: if iron, coal and steel are omitted from the surviving companies, 1924–39, σ^2 rises from 5·18 to 7·74 as compared with the rise of 5·45 to 7·66 given in Table 8. His last question we find a little puzzling: if more "middle-sized" firms appeared we would suppose that concentration had *de*creased.

Mr. Bluett suggests that we should have taken account of changes in the price level. This has, in fact, been done "automatically" in all our measures, since they are based on a logarithmic transformation; a general rise in prices affects the mean (cf. our remarks in section 4.2), but not the variance of the logarithms. Similarly it does not affect the slope of the regressions.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

As a result of the ballot taken during the meeting, the candidates named below were elected Fellows of the Society:

Rex Alan Acton.
Donald Frederick Hurst Bonham.
John Brodrick.
Kum Chee Chan.
Philip Ernest Chavasse.
John White Crofts.
John Barron Eldridge.
Alexander John Gale.
John Clifford Gower.
Robert Allen Hastie.
Albert Thomas Healey.
Cyril Arthur Green.
Edgar Charles Layton.
Brian Davy Lewis.
Ronald Ma.
Ronald Ian Murray Macaulay.
John Corbett McDonald.
Henry Leslie Milnes.
Stanley Francis Mundy Neill.
Matthew Henry Oram.
John Michael Parker.
John Plymen.
Kenneth Arthur Senior Potter.
Andrew Purdon.
Mohammed Ghulam Rabbani.
John Robert Stephen Revell.
Stanley Jack Rose.
John Anthony Rowntree.
Howard Russell Simpson.
Jacob Spector.
John Sidney Smith.
Ernest John Tucker.
Raymond Albert Vandyk.
Bernard Thomas Warner.
Peter Ian Welding.
Edgar Lloyd Williams.
William Henry Rowe Williamson.
Clive Woodcock.

Corporate Representative

Alfred Walter Gent *representing* Standard Telephones and Cables Ltd.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Richard Jones: A Reappraisal

By R. Glenday

The centenary of the death of Richard Jones (1790–1855) who succeeded Malthus as Professor of History and Political Economy at East India College, Haileybury, deserves to be commemorated not only because Jones was one of the founders of the London Statistical Society (now the Royal Statistical Society), but also because he was not simply a "forerunner"—as has been so frequently alleged—but one of the actual progenitors of the inductive approach to economic problems in the nineteenth century. As a thinker he was superior to Malthus.

Fellows of the Society are indebted to L. G. Johnson, lecturer in Economics at Leeds University, for the pamphlet he issued at the close of last year (for private circulation) under the title *Richard Jones Reconsidered.* This suggests that Jones's proper title to "economic fame" is that he was the founder of the English Historical School. Jones's most important work was his *Essay on the Distribution of Wealth and the Sources of Taxation* published towards the end of 1831. The idea that prompted this work was that Adam Smith had done all that could be expected of one man on the production of wealth, but little on the subject of its distribution. Jones decided to make up the deficiency by writing an essay to deal with rent, wages and profit by means of a new method of investigation intended to revolutionize the study of political economy. The only part he found time to complete was that on rent.

Jones's method of approach was inductive, which he described as the "ascending method", by which he meant the method of moving "towards" a conclusion. By contrast the deductive approach, which started from abstract principles, was "downward". By adopting the latter Ricardo had imposed upon his readers a definition of rent which did not fit the facts about rents at different times and places. Jones, on the other hand, aimed at founding his conclusions on a wide observation of contemporary facts, aided by the study of history. The world he set out to study was not an imaginary world inhabited by abstract "economic men", but the real world with the different forms which the ownership and cultivation of land and, in general, conditions of production and distribution, assumed at different times and places. His recognition of such different spheres of life in communities occupying different stages in the progress of civilization led to his proposal for what he called a "Political Economy of Nations". This was a protest against the practice of taking the exceptional state of facts which exists, and is indeed only partially realized, in a small corner of our planet as representing the uniform type of human societies, and ignoring the effects of early history and the special development of each country as influencing economic phenomena.

The difficulty was, with the limited knowledge of customs and practice available in those days, to obtain a sufficiency of facts. As a consequence, Jones's "inductions" were not "complete". They could not have been, of course. "Economists of a hundred years after Jones, provided with abundant data and using techniques of research he could not have dreamed of, do not find it easy to investigate the subject of the 'prosperity of nations' ".

But this was far from being the only—or even the main—obstacle which had to be overcome before the new method could make headway. Jones and his supporters in the induction school found themselves opposed by those—the great majority in those days—who believed with Whateley, Drummond professor of political economy at Oxford, that the "principles of action are known by consciousness and do not require detailed observation".

Few to-day appreciate the grip which the deductive method used to have on social and economic thinking—and indeed still had until very recently, in the field of historical study in general. As a contemporary historian has observed, "The old historians assumed history was 'linear'. They believed fundamentally in progress. They measured civilization by technical advance and liberal institutions and traced its pedigree from Greece through Rome to Modern Europe". The moderns have very different views. "They have jettisoned belief in progress, lost confidence in the uniqueness of Europe. Civilization seems to them not single and linear but multiple and cyclic. There are several civilizations which, being independent of each other, not mere stages in a stream, can be compared". Jones would have felt at home among these "heretics". As it was, by the end of 1832, he knew that his book had failed to open a new phase in the study of political economy. Can one blame this poor "curate of Brasted" suffering from ill-health, when appointed Professor of King's College, London, in the following year, with all the additional work and obligations

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

this entailed, for abandoning all hope of establishing his school of induction and the periodical which was to give it organized form and handing over his "splutting torch to the British Association and the London Statistical Society"?

The first step was the putting forward to the newly-founded British Association of a series of recommendations which were in effect a re-statement of views that had been expressed over a period of several years in correspondence between a group of Cambridge scientists which included the mathematician Airey, the astronomer Herschel, the political economist Jones and the mineralogist Whewell. At the meeting of the association at Cambridge in 1833 the president announced "the somewhat irregular appearance of 'a new bantling' among the sections—one dealing with statistical inquiries".

The new step was the passing by that section in London in February, 1834, of a resolution moved by Malthus and seconded by Jones in the following terms:

> "Following the spirit of the instructions received by the Committee at Cambridge, it is advisable to take immediate steps to establish a Statistical Society in London, the object of which shall be the collection and classification of all facts illustrative of the present condition and prospects of society".

In the same year Jones succeeded Malthus in the chair of History and Political Economy of East India College, where the above resolution provided the background for his subsequent work on inductive thinking until the day of his death in 1855. The Statistical Society itself was inaugurated on March 15th, 1834.

As L. G. Johnson justly sums up, "the founding of an institution of research (the London Statistical Society) has been underestimated because writers by insisting on designating Richard Jones, the economist of the group 'a forerunner' have displaced it from its proper historical setting".

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Wholesale Prices in 1955

By the Editor of "The Statist"

Method of Construction of the Index Numbers

The following table illustrates the method of construction of the index numbers. The index numbers here given are based on the average prices for the eleven years 1867–1877. Take, for instance, the *Gazette* price of English wheat:

	s.	d.			
Average, 1867–77 . .	54	6,	average point.		
„ 1914 . . .	35	0,	64 or 36 per cent.	*below*	the average point.
„ 1930 . . .	80	7, 148	„ 48 „	*above*	„ „
„ 1936 . . .	53	3, 98	„ 2 „	*below*	„ „

The individual index numbers, therefore, represent simple percentages of the average point. The 45 articles are grouped in six categories. They are shown in detail in Table 5.

The general average is drawn from all 45 descriptions, which are treated as of equal value, and is the simple arithmetic mean.

Table 1

Construction of Index Number for 1955

1867–77 = 100

	Number of Commodities in Index	*Total Numbers*		*Index for* 1955
		1867–77	1955	
General average	45	4,500	16,638	370
Food	19	1,900	6,246	329
Vegetable food	8	800	2,553	319
Animal food	7	700	2,570	367
Sugar, coffee and tea . . .	4	400	1,123	281
Materials	26	2,600	10,392	400
Minerals	7	700	3,703	529
Textiles	8	800	3,282	410
Sundry materials . . .	11	1,100	3,407	310

Table 2

*Annual and Monthly Index Numbers**

1867–77 = 100

	Year	*Jan.*	*Feb.*	*Mar.*	*April*	*May*	*June*	*July*	*Aug.*	*Sept.*	*Oct.*	*Nov.*	*Dec.*
1902–1911 .	74	73·9	74·0	74·2	74·5	74·9	74·5	74·5	74·3	74·3	74·5	74·4	74·8
1912–1921 .	148	145·2	146·3	146·9	147·7	147·9	147·2	147·9	149·2	149·0	149·3	149·1	149·2
1922–1931 .	120	122·7	123·0	122·5	122·0	121·0	119·3	119·0	118·0	118·0	118·6	119·2	118·5
1932–1941 .	97	96·0	96·5	96·5	96·8	97·3	96·9	97·5	97·8	98·7	98·8	99·4	101·4
1942–1951 .	231	221·4	224·1	224·6	227·3	228·0	227·1	226·7	227·3	231·2	235·3	238·5	241·2
1943 . .	155	153·3	153·2	154·0	154·9	155·6	155·4	156·0	154·4	154·6	153·7	153·9	153·9
1944 . .	160	154·9	155·1	155·7	157·9	159·5	160·0	161·9	161·5	160·1	159·6	160·3	160·7
1945 . .	164	161·4	161·5	162·1	162·4	164·0	166·7	165·2	163·2	162·7	162·7	162·7	163·2
1946 . .	186	165·6	166·2	166·6	168·4	169·3	169·7	174·4	175·7	180·5	196·2	198·1	200·5
1947 . .	230	206·7	209·6	212·1	215·9	216·5	218·1	223·4	225·9	228·9	236·0	240·8	246·5
1948 . .	260	250·5	253·4	256·0	256·4	260·0	263·9	260·4	260·1	258·1	261·5	263·8	266·0
1949 . .	274	267·6	266·8	265·4	272·0	269·2	267·2	263·3	262·5	285·7	291·5	293·7	297·4
1950 . .	324	300·2	301·0	301·7	305·2	308·3	310·3	317·7	332·1	335·7	349·2	363·1	372·3
1951 . .	402	404·8	420·5	419·3	425·7	420·6	405·2	394·5	387·9	395·9	392·0	396·9	399·7
1952 . .	380	403·5	394·8	384·4	377·5	372·2	379·5	374·4	370·4	369·9	368·9	369·4	372·9
1953 . .	366	374·5	375·2	374·9	370·3	368·1	367·5	361·3	357·9	357·0	357·3	577·3	360·0
1954 . .	361	359·8	361·2	366·1	366·5	363·9	361·2	355·5	353·5	353·7	356·8	362·9	371·7
1955 . .	370	378·0	376·5	375·8	370·2	370·1	370·1	373.1	363.9	366·8	372·4	370·9	374.1

* The average of the twelve monthly figures of each year does not necessarily coincide with the annual figures, as the latter are calculated mostly from the average of 52 weekly quotations, while the former are based on end-of-the-month prices.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 3

Annual and Quarterly Index Numbers for Groups

1867–1877 = 100

	All Groups	Food				Materials			
		Total	Vegetable Food	Animal Food	Sugar, Coffee and tea	Total	Minerals	Textiles	Sundry Materials
1902–1911 . .	74	71	66	88	49	77	90	70	74
1912–1921 . .	148	146	136	167	100	154	166	151	148
1922–1931 . .	120	115	101	150	80	122	137	128	110
1932–1941 . .	97	88	84	116	49	103	131	94	91
1942–1951 . .	231	172	201	169	121	273	300	306	232
1943 . . .	155	138	156	156	72	167	187	166	156
1944 . . .	160	137	152	156	73	178	197	182	161
1945 . . .	164	139	155	156	78	182	209	189	159
1946 . . .	186	140	155	154	88	219	239	231	198
1947 . . .	230	156	191	149	100	284	304	295	263
1948 . . .	260	171	217	155	107	324	368	348	279
1949 . . .	274	201	240	185	151	328	382	356	273
1950 . . .	324	235	268	204	224	390	416	491	299
1951 . . .	402	265	305	227	249	502	516	639	394
1952 . . .	380	294	336	285	227	443	532	475	363
1953 . . .	366	325	365	317	261	399	473	430	329
1954 . . .	361	326	330	325	320	386	476	429	298
1955 . . .	370	329	319	367	281	400	529	410	310
1954 1st Quarter .	362·4	332·1	354·0	316·3	315·8	384.5	462·7	433·6	299·0
2nd „ .	363·9	332·2	345.7	311·2	342·3	386·9	478·6	429·5	295·0
3rd „ .	354·2	317·7	319·5	332·9	307·8	382·0	480·7	422·4	289·9
4th „ .	364·5	326·6	317·6	339·6	321·7	391·1	491·2	432·2	297·5
1955 1st Quarter .	376·8	339·8	333·7	365·8	306·2	407·2	517·1	437·6	307·2
2nd „ .	370.1	330·6	323·9	375·0	266·1	399·1	510·6	417·0	315·1
3rd „ .	367·9	324·2	312·8	383·6	243·2	400·0	542·3	401·3	308·5
4th „ .	372·5	330·8	310·2	389·2	269·7	402·9	558·8	390·6	312·6

TABLE 4

Monthly Index Numbers for Groups

1867–1877 = 100

	Year	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1954													
Food:													
Total . . .	326	328·2	331·6	336·5	334·8	330·7	331·2	329·4	313·9	309.8	314·7	327·8	337·2
Vegetable food .	330	355·9	354·0	352·2	345·1	343·4	348·7	307·0	307·6	302·7	305·6	319·5	327·8
Animal food . .	325	315·4	315·4	318·1	318·1	307·7	307·7	343·2	327·0	328·4	332·9	336·0	350·0
Sugar, tea and coffee .	320	295·0	315·1	337·4	343·6	345·7	337·6	328·6	303·3	291·5	301·2	330·0	333·8
Materials:													
Total . . .	386	382·9	382·9	387·7	389·6	388·0	383·1	377·8	382·5	385·8	387·5	388·6	397·1
Minerals . . .	476	454·1	459·1	474·9	476·5	478·3	480·9	473·4	478·6	490·2	487·7	481·3	504·5
Textiles . . .	429	437·5	430·1	433·3	438·2	430·8	419·5	415·3	424·8	427·2	425·8	433·6	437·1
Sundry . . .	298	297·8	300·1	299·1	291·1	299·5	294·3	289·9	290·6	289·3	295·9	296·9	299·8
1955													
Food:													
Total . . .	329	341·7	340·0	337·7	332·2	329·3	330·2	329·3	315·0	328·3	335·6	328·5	328·2
Vegetable food .	319	335·9	334·0	331·2	320·5	317·2	334·0	327·3	304·1	307·0	314·8	308·4	307·5
Animal food . .	367	360·7	367·7	369·0	371·8	378·9	374·3	385·6	373·5	391·8	390·9	387·7	388·9
Sugar, tea and coffee .	281	320·3	302·5	295·9	286·1	266·8	245·3	234·9	234·6	260·0	280·4	265·1	263·6
Materials:													
Total . . .	400	404·6	403·2	403·7	398·0	400·0	399·3	405·3	399·7	395·0	399·3	401·7	407·7
Minerals . . .	529	515·0	513·8	522·4	509·1	512·0	510·6	542·7	542·0	542·4	546·8	555·5	574·2
Textiles . . .	410	442·9	439·4	430·4	422·2	415·6	413·1	410·9	402·7	390·2	392·6	392·4	386·7
Sundry . . .	310	306·4	306·6	308·7	309·7	317·3	318·4	313·7	307·0	304·7	310·3	310·6	317·0

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

TABLE 5

Average Prices of Each of the 45 Commodities
1951 to 1955

			Average Prices					*Index Numbers* (1866–77 = 100)				
			1951	1952	1953	1954	1955	1951	1952	1953	1954	1955
Vegetable Food												
1	Wheat, English Gazette	*s.* and *d.* per qr.	122 10	129 0	138 0	121 0	98 7	225	237	253	222	181
2	Man. No. 1 Vanc.	,, ,, ,,	106 0	108 11	112 5	124 7	127 6	189	194	200	222	227
3	Flour, G.R.	*s.* per sack (280 lb.)	56⅜	80¾	102⅓	95¼	97¼	123	175	222	207	211
4	Barley, English Gazette	*s.* and *d.* per qr.	124 6	116 4	102 4	91 11	92 11	319	300	275	236	238
5	Oats, ,, ,,	,, ,, ,,	72 11	74 5	68 0	62 10	73 1	280	286	262	242	281
6	Maize, American mixed (*a*)	*s.* per qr.	127⅛	132¼	122⅓	114 0/16	114⅜	391	407	376	353	352
7	Potatoes, good English‡	,, ton	261	295	300	304	374⅝	223	252	257	260	320
8	Rice, Burma. No. 2*	*s.* and *d.* per cwt.	69 0	83 6	107 9	90 0	74 3	690	835	1077	900	743
Animal Food												
9	Beef, prime	*d.* per 8 lb.	125½	158	176	187	234	213	268	300	317	397
10	,, middling	,, ,,	107½	140	150	170	222	215	280	300	340	444
11	Mutton, prime	,, ,,	152½	190	208	196½	214	242	301	330	312	340
12	,, middling	,, ,,	100	111	113	139	182	182	202	206	253	331
13	Pork, large and small, average	,, ,,	152⅜	190	202	196	206	293	365	389	377	396
14	Bacon	*s.* per cwt.	194	281½	339	296	270⅝	262	380	458	400	366
15	Butter	,, ,,	230½	250	292½	346	369½	184	200	234	277	296
Sugar, Coffee and Tea												
16	Sugar, Java floating cargoes (*b*)	,, ,,	45⅜	33	26 5/16	26⅛	26	159	116	94	92	91
17	,, Continental whites	,, ,,	,,	,,	36¼	35 3/16	37 1/16	,,	,,	157	153	161
18	Coffee, Ceylon plantation, low middling (*c*)	,, ,,	520½	471½	448⅜	609	438½	598	542	516	700	504
	Rio, good	,, ,,	395	380	387	481	312	617	594	604	752	488
19	Tea, average import price	*d.* per lb.	39·42	38·36	40·03	54·65	64·61	228	222	234	310	375
Minerals												
20	Iron, Scottish pig	*s.* and *d.* per ton	252·1	291 10	328 9	344 8	372 4	} 367	424	468	500	540
,,	Cleveland (Middlesborough) pig	,, ,, ,,	221·10	255 8	274 10	296 3	324 7					
21	bars, common	£ per ton	21⅞	26 1/16	28⅜	31¼	38⅞	263	325	344	379	471
22	Copper, Standard	,, ,,	220⅝	260	254 1/16	248⅞	351 1/16	294	347	340	332	468
23	Tin, Straits	,, ,,	1,073⅞	964 9/16	721 7/16	707 15/16	740⅜	1,023	919	687	670	705
24	Lead, English pig	,, ,,	163 9/16	135 9/16	91 3/16	96¼	106 1/16	798	661	447	472	520
25	Coal, Best Yorkshire House.†	*s.* per ton	58½	64 1/2	69½	62½	66⅝	266	296	316	284	303
26	average export price	,, ,,	74·88	94·38	89·03	86·75	86·96	599	755	712	694	696

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 5—*cont.*

			Average Prices					*Index Numbers* (1866–77 = 100)				
			1951	1952	1953	1954	1955	1951	1952	1953	1954	1955
Textiles												
27	Cotton, Middling American	*d.* per lb.	46·01	38·92	31·83	32·72	31·70	511	432	354	363	352
28	East Indian	„ „	42·04	36·93	28·73	32·81	31·61	623	547	426	486	468
29	Flax, Medium grade continental retted	£ per ton	486	380½	321½	318⅝	312½	968	737	628	635	592
	Belgian import price	„ „	424·3	312·3	269·9	279·4	243·7					
30	Hemp, Manila for roping	„ „	183 7/12	130 15/16	125	92	92⅜	547	527	474	396	470
	Italian S.B.	„ „	243½	270¼	244¾	216¾	274⅛					
31	Jute, Grade Daisee 2/3	„ „	168⅜	99⅞	79¼	94	84⅝	886	526	417	495	445
32	Wool, Merino, Port Philip, average fleece	*d.* per lb.	177¼	120	126¼	111¼	99⅜	912	621	665	579	493
	Merino, Adelaide, average greasy	„ „	106½	73	80¾	69⅛	54					
33	English Lincoln half hogs	„ „	99	49½	60	63¼	61¾	501	251	304	320	313
34	Silk, Japanese	*s.* per lb.	37⅛	36⅔	40	36⅔	34¼	162	160	174	160	149
Sundry Materials												
35	Hides, River Plate, dry	*d.* per lb.	42¼	25¼	22 15/16	19⅜	17 5/16	558	336	309	281	245
	River Plate, salted	„ „	38 1/16	20 9/16	20¼	19	16 1/16					
	average import price	„ „	47·27	30·96	27·47	25·02	21·71					
36	Leather, dressing hides	„ „	80¾	62½	57	55⅞	49¼	516	388	380	390	387
	average import price	„ „	98 3/16	72½	75¼	79¾	85¼					
	Tallow, Town	*s.* per cwt.	109 5/16	86½	66¾	68¼	73⅜	243	192	148	152	163
	Oil, palm	£ per ton	108⅝	106⅛	83	71⅛	84⅛	278	272	213	183	216
	olive	„ „	387	328¾	305⅜	263 13/16	260	774	657	611	528	520
	linseed	„ „	169	185	180	80¼	102⅝	522	508	443	292	345
	Seeds, linseed	*s.* per qr.	301⅛	272	209¼	182 5/16	207½					
41	Petroleum, motor spirit (c.i.f.)	*d.* per imp. gall.	12·66	13·30	12·32	12·08	11·53	80	84	77	76	73
	Kerosene (burning oil) c.i.f.	„ „ „	11·41	12·58	11·36	10·87	10·86	181	199	180	172	172
	gas oil, c.i.f.	„ „ „	11·19	11·84	10·76	10·32	10·70	281	297	270	259	269
42	Soda, crystals	*s.* per ton	125	190	190	190	212½	136	206	206	206	231
43	Sodium nitrate	*s.* per cwt.	27⅜	30	29¾	27	27¼	195	214	212	193	195
44	Indigo, Bengal, good consuming	*s.* per lb.	9⅛	14	12½	12½	12½	126	193	172	172	172
45	Timber, hewn, average import price	*s.* per load	367·83	376·00	374·64	338·12	338·44	806	829	758	717	762
	sawn, or split, average import price	„ „	550·31	569·24	489·49	479·04	530·16					

(*a*) La Plata from 1924–40; Argentine (Feeding Stuffs) from 1941–52; La Plata subsequently. (*b*) Raw centrifugals. 96 per cent. Pol., from 1924. (*c*) East India, good middling from 1908–1947; Kenya from 1948. * During war period to mid-July, 1953, quotations represent any description of whole rice under the Rice Order. † Owing to change in selling methods the quotation for house coal alone, no longer ascertainable; from 1954 the quotations relate to the average pit price of all grades of Yorkshire coal. ‡Average of weekly quotations during the eight months, January to April and September to December.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 6

Index of Silver Prices

The base of the Index Numbers given below is 60·84*d.* per standard ounce = 100, this being a parity of 1 fine oz. of gold to $15\frac{1}{2}$ standard ozs. of silver.

	Price per oz. Standard (d.)		*Index Number*	
	Average for year	*At end of December*	*Year*	*At end of December*
1942	$23\frac{1}{2}$	$23\frac{1}{2}$	18·1	18·1
1943	$23\frac{1}{2}$	$23\frac{1}{2}$	18·1	18·1
1944	$23\frac{1}{2}$	$23\frac{1}{2}$	18·1	18·1
1945	$30\frac{1}{2}$	44	23·1	31·1
1946	$48\frac{5}{8}$	$55\frac{1}{2}$	36·7	41·6
1947	$44\frac{7}{16}$	45	33·3	33·8
1948	45	$42\frac{1}{2}$	34·1	31·9
1949	$49\frac{1}{9}$	64	32·8	33·3
1950	$64\frac{4}{5}$	70	33·7	36·4
1951	$77\frac{27}{32}$	77	40·5	40·1
1952	$74\frac{3}{8}$	$72\frac{3}{4}$	38·7	37·9
1953	$73\frac{15}{16}$	$73\frac{3}{4}$	38·5	38·5
1954	$73\frac{1}{2}$	$74\frac{1}{4}$	38·4	38·6
1955	$77\frac{1}{2}$	$78\frac{3}{4}$	40.0	40·8

TABLE 7

Gold and Silver: Production

	Gold Value of output (£ *thousand*)	*Silver* (*million ounces*)				
		United States	*Mexico*	*Canada*	*Australia*	*Central and South America*
1942	289,153	54·5	84·9	21·8	14·2	31·8
1943	226,397	44·8	71·2	18·6	10·3	29·3
1944	209,031	37·4	63·0	14·8	9·4	29·7
1945	199,116	29·3	61·1	14·0	8·1	27·4
1946	202,435	21·4	48·3	13·6	9·0	26·1
1947	206,434	36·1	49·2	13·5	9·5	23·6
1948	214,647	36·1	45·8	17·0	10·1	23·0
1949	320,887	34·6	49·5	17·6	9·8	24·0
1950	329,580	42·0	49·0	23·2	10·7	28·0
1951	322,623	39·9	43·8	23·1	10·8	28·7
1952	329,381	39·8	50·4	25·2	11·4	32·6
1953	325,500	37·7	47·9	28·3	12·5	34·1
1954	344,865	37·1	39·9	30·7	11·5	32·4
1955*	340,732†	36·0	45·0	27·7	11·0	31·0

* Provisional. † Excluding U.S.S.R.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Sources and Nature of Statistical Information in Special Fields of Statistics

Banking Statistics

By W. Manning Dacey*

In the post-war years the publication of the *Statistical Digest* has made easy the path of the general user of banking, as of so many other, statistics. For this period the tables in the "Finance" section present a convenient summary of the weekly Bank of England returns, the monthly returns published by the Committee of London Clearing Bankers, the quarterly analyses of bank advances published by the British Bankers' Association and much other useful financial information. In addition, the *Digest* is now the medium of publication of the analysis of notes and coin outstanding and of the average estimated circulation with the public prepared by the Bank of England and previously published in the Bank's own statistical summary.

Nowadays it is only the more specialized student, requiring figures as soon as they are available to the public or wishing to extend his researches to years before 1945, who requires to go beyond the *Digest* to the original sources, or even to the financial press, for his information. These notes are therefore designed chiefly to provide some indication of the continuity of the available material and some guidance on points of interpretation.

Bank of England

(1) *Weekly Return*

One of the schedules to the 1844 Bank Charter Act specified the form of the Bank's weekly Return, as follows:

Issue Department

Notes Issued	Government Debt Other Securities Gold Coin and Bullion Silver Bullion

Banking Department

Proprietors' Capital Rest Public Deposits (to include Exchequer, Savings Banks, Commissioners of National Debt and Dividend Accounts)	Government Securities (including Dead Weight Annuity) Other Securities Notes Gold and Silver Coin

From 1844 to 1928 the Return was published in this form without changes of any kind other than the dropping of the item "Silver bullion" by the Issue Department after 1890. Subsequent modifications have left the form of the Return basically unchanged, while increasing its usefulness. After the 1928 Act, which merged the Treasury note issue with the Bank's own issue, the *Issue Department* divided "Notes issued" into notes "In circulation" and "In Banking Department" respectively. On the assets side the item "Other securities" was sub-divided into "Other Government securities" and "Other securities", while the "Amount of fiduciary issue" was shown as the total of Government debt, the two securities items and—a new item—"Silver coin". (In January, 1946, the latter item was renamed "Coin other than gold".)

Notwithstanding the depreciation of sterling against gold after 1931, the Bank of England continued to value its gold coin and bullion at the old statutory price until the passage of the Currency and Bank Notes Act of 1939; since then a valuation price, related to the current market

* With the assistance of Mrs. R. Kirby-Johnson and other members of the staff of Lloyds Bank.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

price for the metal and expressed in shillings and pence per *fine* ounce, has been indicated in the Issue Department's return.

The *Banking Department* in 1928 split the item "Other deposits" into "Bankers' deposits" and "Other accounts", while on the assets side "Other securities" were sub-divided into "Discounts and Advances" and "Securities", both changes permitting rather greater insight into movements affecting the basis of credit. It was in 1928 also that the small item "seven-day and other bills"—a relic of earlier days—disappeared. In August, 1948, the item "Public deposits" was also sub-divided, the normal components of that item being entitled "Public accounts" while a new item "H.M. Treasury Special Account" made its appearance; this was part of the curious, and from the economic viewpoint meaningless, book-keeping called for by the United States legislation in respect of Marshall Aid.

(2) *Statistical Summary*

Following a recommendation of the Macmillan Committee, the Statistical Summary previously compiled by the Bank of England for internal use and private circulation was issued to the public for the first time in January, 1932, and it appeared monthly until superseded by the *Statistical Digest* at the beginning of 1946. In addition to providing a useful compilation of financial information generally, the Summary was itself the original medium of publication of various data, notably:

(a) *Estimates of currency circulation in the hands of the public.* The issue of May, 1932, contained monthly estimates of notes and currency in the hands of the public for the period January, 1921, to April, 1932, inclusive, together with two indices each having 1924 as the base year, one seasonally adjusted and the other not. Both monthly and annual figures for the period 1921–1938 were published in the issue of January, 1939, together with a seasonally adjusted index revised and with the base year changed to 1935. Three years later a fresh series was issued covering the period January, 1935, to January, 1942, the index being no longer seasonally adjusted; thereafter figures on this basis appeared monthly until publication of the Summary was suspended at the end of 1945.

To complete the record it should be mentioned that the first issue of the *Statistical Digest* gave annual figures beginning 1935 but with the year 1938 as the basis of the new index, and this series has continued monthly up to the present time. Linking the various series together, and using the later figures in each case for overlapping periods, gives the following series (1938 = 100) for the years 1921 to 1955 inclusive:

TABLE 1

Estimated Notes and Currency in Hands of Public

1921	85·9	1930	72·6	1940	117·7	1950	281·4
1922	78·2	1931	73·3	1941	136·3	1951	292·1
1923	75·6	1932	74·8	1942	172·4	1952	310·0
1924	76·3	1933	76·1	1943	209·2	1953	330·8
1925	75·0	1934*	77·5	1944	247·8	1954	350·9
1926	74·2	1935	80·6	1945	283·2	1955	374.9
1927	73·5	1936	88·5	1946	300·4		
1928	73·3	1937	98·9	1947	305·2		
1929	73·1	1938	100·0	1948	277·8		
		1939	102·9	1949	279·8		

* 1921–34 ten banks, thereafter eleven (the figure for 1935 being adjusted to include the District Bank).

(b) *Note issue by denominations.* A consecutive series of annual figures (representing the average of Wednesdays) for the period 1921 to 1945 inclusive is obtainable from the issues of the Summary for January, 1935, July, 1939, and December, 1945.

A similar analysis (though relating only to the end of February in each year) has been published, for the years 1939, 1943 and thereafter consecutively to 1954, in the annual reports of the Bank of England.

(c) *Division of clearing bank deposits between current and deposit accounts.* The monthly clearing bank figures contained in Appendix I of the Macmillan Report showed, for the period January, 1919, to March, 1931, inclusive, the distribution of total clearing bank deposits between current accounts, on the one hand, and "deposit and other accounts", on the other hand. The

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

annual averages of these figures were reproduced in the July, 1932, issue of the Summary, together with a series of monthly figures commencing in January, 1931. This latter series was continued in the issue of February, 1933, and from April, 1933, onward appeared monthly until the suspension of publication. In these figures, offices outside England and Wales were excluded up to December, 1937.

For the post-war period the series has been continued in the monthly *Digest*, though since December, 1947, the division between current and deposit accounts has been announced in the monthly clearing bank return itself. From Table 2, it will be seen that between 1932 and 1942 the proportion of current accounts rose from about half to two-thirds of the total. This was undoubtedly a reflection of the very low rates of deposit interest ruling during the "ultra-cheap" money period of a 2 per cent. Bank Rate, which meant that account-holders lacked the normal incentive to transfer to deposit account any balances surplus to their probable requirements. There has been a slight reversal of this movement since flexibility was restored to Bank Rate in 1951.

TABLE 2

Current Accounts as a Percentage of Total Clearing Bank Deposits
(*Annual Averages*)

	%		%		%		%
1919*†	67·2	1930	52·9	1940	59·3	1950	66·2
1920	63·8	1931†	52·7	1941	63·5	1951	66·5
1921	58·4	1932	49·5	1942	65·6	1952	65·1
1922	58·5	1933	51·3	1943	66·8	1953	63·9
1923	59·8	1934	51·8	1944	66·6	1954	63·8
1924	58·8	1935*	53·8	1945	66·6	1955	63·7
1925	57·6	1936	54·9	1946	66·3		
1926	57·3	1937	55·0	1947	65·3		
1927	56·2	1938	54·6	1948	65·1		
1928	55·8	1939	55·7	1949	66·0		
1929	54·1						

* 1919–35 relate to ten clearing banks, thereafter eleven.
† 1919–30 taken from Macmillan Report; figures for National Bank relate to offices in England only.

(3) *Annual Reports*

Reference has already been made to the analysis of the note circulation by denominations appearing in the annual reports. The latter are also the original source of figures (for 1939, 1944 and thereafter consecutively) of notes issued and paid during the year—Table 3—and (for 1948 onwards) of sterling transfers between countries.

In connection with the annual report, studies of U.K. overseas investments have been issued annually since 1950, providing a continuation of the pre-war series of studies published under the name of Lord Kindersley in the *Economic Journal*.

TABLE 3

Bank of England Notes

Year ended February	*Issued* (£*m*)	*Paid* (£*m*)	*In Circulation at End of Year* (£*m*)
1939	969	968	478
1944	831	664	1,086
1945	857	726	1,217
1946	935	830	1,322
1947	927	872	1,377
1948	961	1,105	1,233
1949	1,110	1,111	1,232
1950	815	797	1,250
1951	916	877	1,289
1952	1,095	1,015	1,369
1953	1,136	1,035	1,470
1954	1,144	1,064	1,550

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Committee of London Clearing Bankers

Monthly Clearing Bank Returns

It was in April, 1921, that a monthly statement for the London clearing banks (then ten in number) was first published by the Committee, though *The Economist* had compiled monthly figures for many years prior to July, 1915 (when publication was suspended owing to war conditions), and resumed publication in January, 1921.

The continuity of the figures was interrupted by the District Bank becoming a clearing bank in January, 1936, so that strict comparison with earlier periods would necessitate the use of a series covering only the original ten banks. In practice, this adjustment is probably no longer worth making, as the effect is slight and there have been other breaks in the continuity of the figures. The figures represented the average of returns drawn up weekly (though on varying days for the different banks) until September, 1939, when the weekly averages were replaced by figures relating to a single day in the second half of the month, each bank choosing a date convenient to itself. On and after December 31st, 1946, the figures have been aggregated on a date common to all the banks, "synchronizing with the weekly return of the Bank of England and compiled on the third Wednesday in each month except in the months of June and December", when the statements are made up on the 30th and 31st respectively. This was, of course, a change of considerably more than statistical interest, since it signalized the abandonment of the former practice of "window-dressing" and the decision to aim consistently, and not merely on accounting dates, at the maintenance of a cash ratio of around 8 per cent.

On the assets side, Treasury deposit receipts appeared as a new item in the return for July, 1940, and an actual holding of Treasury deposit receipts was last shown in that for February, 1952, though a blank line for that item continues to be included in the returns. Since November, 1951, "Bills discounted" has been divided into "Treasury bills discounted" and "Other bills discounted". As already noted, the division between current and deposit accounts has been stated since December, 1947, as a footnote to the return. This has also included since August, 1952, an additional two columns showing the changes on the month and on the year in all the aggregate figures.

For most purposes, the published totals of clearing bank deposits are regarded as equivalent to the total liquid assets of the public in the form of bank credit, though in principle this is subject to certain qualifications. In the first place the figures necessarily take no account of outstanding overdraft limits; and the additional liquidity which these confer (reducing the need to hold actual bank balances) may vary from time to time, so that even a "purified" figure of the banks' liabilities to the public would not represent an accurate measure of the liquid position of the public. Furthermore, the figures of Current, Deposit and Other Accounts exceed the banks' liabilities to the public by the inclusion of:

(*a*) The internal accounts of the banks themselves.
(*b*) The counterpart of cheques in course of collection drawn on other banks; and
(*c*) The counterpart of "items in transit", which may perhaps best be described in the "internal" equivalent of cheques passing through the clearings, representing typically cheques, in process of collection for customers, drawn by customers of other offices of the same bank.

Items (*b*) and (*c*) have a pronounced seasonal movement, rising sharply in June and December especially and correspondingly inflating the figures of "gross" deposits. For many purposes, therefore, a figure of "net" deposits is more useful, as the monthly changes in "net" deposits correspond more closely to changes in the banks' liabilities to the public and to changes in the combined totals of the principal assets (cash, call money, discounts, investment and advances) "Net" deposits are usually defined as gross deposits less the item "Balances with and cheques in course of collection on other Banks in the U.K. and Republic of Ireland".* Since "Items in transit" are unfortunately now shown as a separate asset only by Lloyds Bank,† it would appear

* This is, for example, the treatment adopted in the compilation of Lloyds Bank's seasonally adjusted index of net deposits (Table 4) though in the statistics compiled by the London and Cambridge Economic Service "Items in transit" (of the one bank in question) are also deducted.

† Martin's Bank published such an item from January, 1930, to August, 1943, and it was revealed on one occasion (April, 1921) by Williams Deacon's.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that most of the other banks include "Items in transit" in their advances item, with the effect that for this among other reasons month to month changes in the advances total do not necessarily reflect equivalent changes in lendings to the domestic public.

Bank deposits show a regular and pronounced seasonal variation dominated by the heavy influx of revenue to the Exchequer (and accompanying retirement of floating debt) during the first quarter of the calendar year and the normal seasonal deficit, financed by expansion of the floating debt, during the remaining nine months. Allowance for this, and for the seasonal variations in transit items, are the chief points to be borne in mind when interpreting changes in the figures over short periods. An index of net deposits, adjusted for seasonal variations, is prepared by Lloyds Bank and released to the Press to coincide with the issue of the monthly clearing bank statements.

TABLE 4

Lloyds Bank Index of Net Deposits (1948 = 100) *Adjusted for Seasonal Variations*

	1948	*1949*	*1950*	*1951*	*1952*	*1953*	*1954*	*1955*	*1956*
Jan.	95·7	100·5	101·1	103·6	102·3	103·8	107·0	111·1	106·3
Feb.	97·8	100·8	101·1	104·5	102·0	104·2	107·6	112·0	105·6
Mar.	101·0	101·2	100·6	104·5	102·1	104·9	108·0	110·3	
Apl.	100·9	100·8	100·7	105·3	102·1	104·0	108·8	109·8	
May	100·0	100·8	100·1	103·7	101·7	104·7	109·0	109·0	
June	100·5	101·4	100·6	103·7	101·9	105·5	109·2	108·1	
July	100·2	101·3	101·1	104·3	102·3	105·3	109·3	107·8	
Aug.	100·7	100·9	101·5	103·8	102·8	105·8	110·0	107·8	
Sept.	100·7	101·1	102·2	103·6	103·3	106·3	109·8	106·5	
Oct.	101·1	101·4	103·7	103·3	102·6	106·3	110·1	105·8	
Nov.	100·9	101·0	103·9	102·9	103·0	106·8	110·9	105·4	
Dec.	100·4	101·3	103·6	102·5	103·8	107·4	110·6	105·6	

In longer-term comparisons, the chief break in continuity (other than the inclusion of the District Bank in 1936) arises from the abandonment of "window-dressing" at the end of 1946. Before that date, the collective cash ratio dipped below 10 per cent. only on one occasion (October, 1936), though a fallacy of composition was involved in aggregating returns prepared on varying days. Some estimate of the amount of window-dressing can be made by comparing on the one hand the cash holdings shown in the clearing bank returns and on the other hand the Bank of England figures for clearing bank till money and "bankers deposits" in the Banking Department. The results can be no more than tentative guesses, since the dates do not coincide (the "till money" figures, for example, relate to the *average* of Wednesdays) and the proportion of non-clearing bank funds in "Bankers' deposits" is not known and presumably fluctuates, though usually assumed to be about 5 per cent. A point frequently overlooked is that, to the extent that the pre-1947 clearing bank returns conveyed an exaggerated impression of the banks' "true" cash holdings, they must correspondingly have understated "true" holdings of money market assets.

British Bankers' Association

Classification of Bank Advances.

One of the many statistical recommendations of the Macmillan Committee was that the banks should publish annually, "or at quarterly intervals if that be practicable", a classification of their loans and overdrafts by type of borrower. The Macmillan Report itself published an analysis, in thirteen groups, of the advances of ten London clearing banks on various dates in late 1929 and early 1930. Three similar classifications (under fourteen heads and covering all eleven banks) were published in the Bank of England Statistical Summary, relating to various dates in the latter half of 1936, 1937 and 1938 respectively.

Since February, 1946, a quarterly analysis has been issued by the British Bankers' Association in which the advances within Great Britain of the constituent banks were sub-divided into twenty-five groups until February, 1955, when the "Other Financial" category was split to show separately loans to hire-purchase finance companies.

The largest single group is still the "Personal and Professional" category, though this has declined, under the influence of official "directives", from nearly 30 per cent. of the total just

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

after the war to considerably less than 20 per cent. in the latest return. This group, incidentally, includes in addition to personal and professional loans in the obvious sense advances to executors and trustees for payment of death duties and other purposes, as well as advances for the purchase of house property to be occupied by the owner-borrower.

For many years past, the five groups separately enumerated in Table 5 have made up about two-thirds of the total. A category of increasing importance since the war has been loans to "Public Utilities other than Transport", which rose to more than 9 per cent. of the total in mid-1955 but has since fallen back sharply. This group comprises, in addition to water companies, the nationalized gas and electricity industries, whose practice is to finance their current requirements by means of bank finance, funding this indebtedness by the issue of long-term loans on the capital market at convenient intervals.

TABLE 5

Classification of Bank Advances

Yearly Average	*Personal and Professional*	*Agric. and Fishing*	*Other Financial*	*Retail Trade*	*Food, Drink and Tobacco*	*Engineering*	*All Advances*
1946 £m	261·4	73·7	77·1	68·4	59·4	62·5	896·2
%	*29·2*	*8·2*	*8·6*	*7·6*	*6·6*	*7·0*	
1947 £m	321·1	89·0	97·2	102·3	75·1	82·3	1,128·6
%	*28·4*	*7·9*	*8·6*	*9·1*	*6·7*	*7·3*	
1948 £m	365·9	115·4	103·5	138·9	81·7	112·4	1,354·6
%	*27·0*	*8·5*	*7·6*	*10·3*	*6·0*	*8·3*	
1949 £m	400·0	137·2	125·9	159·2	101·3	109·6	1,494·6
%	*26·8*	*9·2*	*8·4*	*10·6*	*6·8*	*7·3*	
1950 £m	419·4	163·6	159·5	180·1	121·0	93·1	1,683·9
%	*24·9*	*9·7*	*9·5*	*10·7*	*7·2*	*5·5*	
1951 £m	434·3	190·4	200·2	207·1	142·7	109·5	1,920·6
%	*22·6*	*9·9*	*10·4*	*10·8*	*7·4*	*5·7*	
1952 £m	402·2	198·9	177·6	189·4	158·1	150·0	1,930·6
%	*20·8*	*10·3*	*9·2*	*9·8*	*8·2*	*7·8*	
1953 £m	366·7	200·3	179·1	169·6	132·5	153·6	1,815·8
%	*20·2*	*11·0*	*9·9*	*9·3*	*7·3*	*8·5*	
1954 £m	374·3	213·7	193·1*	176·9	133·6	137·7	1,904·1
%	*19·7*	*11·2*	*10·2*	*9·3*	*7·0*	*7·2*	
1955 £m	389·1	233·6	215·6*	192·7	162·5	158·7	2,120·2
%	*18·4*	*11·0*	*10·2*	*9·1*	*7·7*	*7·5*	
1956 Feb. £m	347·9	229·1	202·8*	185·7	163·1	180·1	1,961·1
%	*17·7*	*11·7*	*10·3*	*9·5*	*8·3*	*9·2*	

* Includes the following amounts in respect of hire purchase finance companies:

1954	£m	%	1955	£m	%	1956 Feb.	£m	%
	15·1	*·8*		38·6	*1·8*		35·4	*1·8*

Bankers' Clearing House

(1) *Annual Reports*

While figures of bank clearings have been made public for the better part of a century, there have been many significant changes both of basis and of mode of presentation.

The annual report for 1901 forms a convenient point of departure. It consists of three documents, each covering the period 1868–1901 inclusive, as follows:

(i) Annual totals (a) of aggregate clearings, and (b) of clearings on the fourth day of each month (the actual date of payment of bills of exchange nominally falling due on the first of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

month), on Consols settling days, on Stock Exchange account days and on days following the fortnightly settling days.

(ii) Daily averages for each year on "ordinary" days and on the various kinds of special days mentioned above. The object of this sub-division is clearly to segregate as far as possible the influence on the figures of large financial transactions.

(iii) A record of monthly, half-yearly and yearly totals for the period.

In the following year, Town and Country clearings are shown separately and in 1908 the Metropolitan clearings make their appearance. It was not until 1923 that separate figures of clearings on the fourth day of the month and on Consols settling days were dropped, their place being taken by the incorporation in the Report for the first time of the Provincial clearings. These were then ten in number, Nottingham and Hull having been added since the turn of the century, while Bradford was first included in 1926 (with figures back to 1911) and Southampton in 1946, with figures back to 1940. The continuity of the Provincial series is greatly affected by numerous changes of area, indicated as footnotes in the returns for recent years.

At the outbreak of war in 1939 a number of changes took place, the combined effect of which is to render extremely hazardous any comparison of pre-war figures with those for subsequent years. The central clearing itself was removed from London to Stoke-on-Trent, and the former Metropolitan and Country clearings, together with part of the Town clearing, were swept into a new General clearing. (In September, 1940, the Town clearing was further reduced to about twenty branches, compared with about 130 branches before the war.) In addition, some changes were made in the categories of items passing through the clearings, while Stock Exchange settlements were suspended and for the first time the clearings were used for items drawn on the Bank of England, formerly a member only on the "outward" side.

Publication of the annual reports continued throughout the war, but although, surprisingly, the Provincial totals continued to be published, the central totals were not at the time sub-divided into Town and General. It was not until 1946 that the relevant figures for the years 1939–1945 were made known.

In the immediate post-war years, the Annual Report was issued in much the same form as before 1939, each consisting of three separate returns and presenting figures for a considerable span of years. An interesting new item of information was added in 1948 with the publication for the first time of the *number*, as distinct from value, of items cleared, annual totals being given for the years 1938, 1947 and 1948. This series (which incidentally makes it possible to calculate the average size of the item cleared) has been continued in subsequent years. The steady growth in the average size of item provides an interesting sidelight on the progress of inflation.

TABLE 6

	Number of Articles Cleared (millions)	*Average Size of Article* £
1938	260	152
1947	289	254
1948	303	265
1949	312	276
1950	321	294
1951	333	327
1952	335	330
1953	345	358
1954	354	408
1955	368	415

The 1948 Report gives a complete run of figures going back to 1910. A radical change in the form of the Report was, however, made in 1949. Since that year it has contained only: (1) quarterly and annual totals of aggregate, Town and General clearings for the latest and for the two preceding years, an indication being given of the number of working days in each quarter; (2) similar figures for the aggregate of the Provincial clearings, with annual figures only for the individual centres.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Monetary totals of the clearings are dominated by the very large financial payments (e.g. for Treasury bills and settlement of Stock Exchange transactions) passing through the Town clearing, the figures of which are three to four times as large as those of the General clearing, even though only about 80 offices are now included, compared with the pre-war 130. Variations in aggregate clearing totals are thus likely to be associated as much with variations in Stock Exchange activity as with variations in the level of national income. The General and Provincial clearings are, of course, more closely related to income-generating payments, but long-term comparisons (of the post-war General clearing with the pre-war Metropolitan and Country clearings) are vitiated by the changes, already mentioned, in the coverage of the Town clearing. In the following table therefore, comparisons of the volume of bank credit and of national income are made with the aggregate central clearings:

TABLE 7

Clearings, Deposits and Income

		Clearings as a Multiple of					
		Clearing Bank Deposits		*Current Accounts only*			
	*London Clearings** (£m)	10 *banks*	11 *banks*	10 *banks*	11 *banks*	*National Income*	*Gross National Product*
1921	34,931	19·3	—	34·1	—	7·8	—
1922	37,161	20·9	—	37·2	—	9·6	—
1923	36,628	21·9	—	37·9	—	9·5	—
1924	39,533	23·7	—	41·6	—	10·1	—
1925	40,437	24·3	—	43·8	—	10·2	—
1926	39,825	23·9	—	43·2	—	10·2	—
1927	41,551	24·3	—	44·6	—	10·0	—
1928	44,205	25·0	—	46·3	—	10·6	—
1929	44,897	24·9	—	47·8	—	10·7	—
1930	43,558	24·2	—	47·3	—	11·0	—
1931	36,236	20·6	—	40·5	—	9·9	—
1932	32,112	17·9	—	37·0	—	9·0	—
1933	32,138	16·5	—	32·9	—	8·6	—
1934	35,484	18·9	—	37·2	—	9·1	—
1935	37,560	18·8	—	35·6	—	9·1	—
1936	40,617	19·0	18·3	35·1†	33·9	9·3	—
1937	42,686	—	18·7	—	34·6	9·2	—
1938	39,610	—	17·4	—	31·8	8·5	7·7
1947	73,330	—	13·0	—	19·9	8·4	7·8
1948	80,210	—	13·6	—	20·8	8·3	7·7
1949	86,060	—	14·4	—	21·8	8·4	7·8
1950	94,241	—	15·7	—	23·7	8·8	8·1
1951	108,773	—	17·7	—	26·6	9·2	8·5
1952	110,584	—	18·2	—	27·9	8·5	8·0
1953	123,373	—	19·7	—	30·9	—	8·3
1954	144,288	—	22·2	—	34·8	—	9·2
1955	152,613	—	23·6	—	37·1	—	n.a.

* 1921–38 Town, Metropolitan and Country clearings; 1947 to date Town and General clearings.
† Assuming proportion of current accounts the same for the District Bank as for the original ten banks.

(2) *Current Releases*

Prior to 1939 daily figures of the central clearings and weekly figures of the Provincial clearings were published each week to coincide with the appearance of the Bank Return. Publication was suspended during the war but resumed on the same basis in 1945. Since January, 1950, however, only monthly figures have been published. The returns for the central clearings show, for the current month and the corresponding month of the preceding year, the figures of Total, Town and General clearings, the number of working days being indicated, with similar comparisons of the cumulative totals for the year to date and the percentage increase or decrease on a working day basis. Similarly, for the Provincial clearings, the returns show monthly totals, with comparative figures for the corresponding month a year earlier, both for the individual centres and

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

for the combined total, together with an indication of the percentage change (on a working day basis) in these figures and in the cumulative total for the year.

Central Statistical Office

Personal Deposits

In four issues of the National Income White Paper (Cmd. 6438, 6520, 6623 and 6784) tables were published showing the proportion of clearing bank deposits, net of advances, belonging to personal account-holders at end-June and end-December dates from June, 1940, onwards. The series was extended to December, 1947, in a table appearing in Statistical Abstract No. 85. In their memorandum to the Royal Comission on Taxation of July, 1951, the British Bankers' Association added figures for 1948 and 1950, while figures relating to June, 1951, were given in oral evidence to the Commission. The returns for December, 1949, 1952 and 1953 can be deduced from a table appended to the paper read to the Manchester Statistical Society by Mr. C. T. Saunders in November, 1954, and showing the changes in net personal bank deposits during the calendar years 1950–1953 inclusive.

The complete series is summarized in Table 8 below. Exceptional care is needed in any use made of these figures, owing to differences of definition. In collecting the figures, the banks seek to include only those accounts conducted by individuals in their private and personal capacity. The range of accounts covered is thus a narrower one than that corresponding to the "Personal and Professional" category in the quarterly classification of advances; and it is, *a fortiori*, narrower than that corresponding to the "personal sector" in the National Income estimates, which includes farmers, sole traders and non-incorporated business generally.

TABLE 8

Personal Share of Net Bank Deposits*

	Total Deposits Net of Advances (£m)	*Personal Deposits Net of Advances* (£m)	*Personal as % of Total* (%)
June Returns:			
1940	1,507	617	40·9
1941	2,087	661	31·7
1942	2,463	704	28·6
1943	2,886	859	29·8
1944	3,330	1,032	31·0
1945	3,990	1,260	31·6
1946	4,162	1,406	33·8
1947	4,539	1,593	35·1
1948	4,614	1,711	37·1
1949	4,577	n.a.	n.a.
1950	4,352	1,693	38·9
1951	4,285	1,676	39·1
December Returns:			
1941	2,522	692	27·4
1942	2,856	820	28·7
1943	3,289	972	29·6
1944	3,791	1,158	30·6
1945	4,035	1,339	33·2
1946	4,705	1,540	32·7
1947	4,728	1,692	35·8
1948	4,822	1,754	36·4
1949	4,679	1,728	36·9
1950	4,724	1,714	36·3
1951	4,402	1,688	38·3
1952	4,711	1,748	37·1
1953	4,987	1,830	36·7

* i.e. gross deposits less advances.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Secondary Sources

It has already been noted that *The Economist* was in the habit of compiling monthly figures of clearing bank returns for some years before these were issued by the Committee of London Clearing Bankers itself. Perhaps the longest continuous series in the banking field was that of the deposits of all joint stock banks published in the periodical Banking Supplements of *The Economist* from 1877 onwards. From the outset, the balance sheets of all the Scottish banks were included, but for

TABLE 9.—*Scottish Banks*

	Deposits	*Investments*	*Advances and Discounts*	*Cash and Money at Call*
1900	107·2	33·4	70·4	26·4
1901	107·3	33·2	70·2	26·9
1902	106·9	33·0	71·5	26·0
1903	104·3	31·2	71·4	24·8
1904	101·9	31·5	69·4	23·9
1905	100·6	31·4	68·7	23·9
1906	105·2	31·2	71·2	26·1
1907	108·7	32·0	73·0	27·0
1908	105·8	33·5	69·5	25·7
1909	106·7	34·5	69·4	25·7
1910	106·7	33·3	70·7	25·6
1911	112·2	33·9	71·6	30·1
1912	118·8	36·6	72·4	33·9
1913	125·9	38·0	77·8	34·0
1914	132·5	44·8	78·4	34·6
1915	140·6	67·0	64·6	36·7
1916	166·6	69·7	85·1	42·0
1917	196·5	71·6	112·2	47·8
1918	222·1	93·9	114·0	57·3
1919	259·6	112·0	130·0	65·6
1920	279·2	103·0	153·1	73·0
1921	291·1	114·5	150·9	70·2
1922	267·8	126·0	122·4	63·8
1923	247·2	112·1	114·3	64·0
1924	242·6	103·9	123·0	61·3
1925	240·9	95·2	131·3	60·3
1926	238·8	91·2	133·1	59·6
1927	242·0	88·6	138·7	61·2
1928	248·8	83·4	149·6	63·3
1929	250·9	83·7	154·1	62·3
1930	256·1	98·1	146·1	63·8
1931	243·6	101·4	134·0	60·4
1932	276·5	136·7	122·9	68·1
1933	281·7	157·7	108·3	66·6
1934	286·7	161·1	106·3	70·5
1935	308·5	177·7	107·1	75·9
1936	317·3	186·3	105·3	76·0
1937	330·5	187·5	114·7	83·3
1938	325·3	184·3	113·7	80·0
1939	331·1	181·4	119·5	88·9
1940	364·4	196·9	127·1	104·2
1941	417·3	241·7	134·5	113·5
1942	453·1	264·3	141·9	130·4
1943	493·8	296·6	143·2	146·6
1944	542·5	316·2	164·3	165·7
1945	590·4	347·0	174·0	170·5
1946	657·8	383·8	186·0	196·7
1947	703·6	407·9	195·3	203·3
1948	750·3	413·3	213·8	209·3
1949	742·4	407·4	211·2	220·1
1950	760·4	403·8	226·6	228·4
1951	751·7	354·4	260·2	240·4
1952	764·4	360·8	242·0	273·2
1953	784·3	396·8	230·7	274·9
1954	834·0	405·8	257·3	296·9

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

some years the coverage was very incomplete not only for the Irish banks but for those of England and Wales. During the 'eighties the omissions were progressively rectified and from 1889 the figures are comprehensive.

Even after this date, the figures of assets are rather less useful, cash often being lumped together with call money or discounts, money at short notice with investments and discounts with advances; as late as 1938 one or two of the smaller banks were still grouping call money with cash and advances with discounts. However, the Banking Supplement itself is now defunct, having expired in June, 1953. While very similar information has been available for many years in the banking supplements of *The Statist*, the figures have not hitherto been aggregated, but it is understood that this may be done in future.

Unlike the London clearing banks, the Scottish banks do not publish monthly statements and their annual balance sheets are drawn up on various dates throughout the year. In Table 9, the figures for 1900–1947 inclusive are based on *The Economist* banking supplements and thereafter on those of *The Statist*.

The bulletins of the *London and Cambridge Economic Service*, first issued January, 1923, published monthly figures of the gross deposits of nine London clearing banks (i.e. excluding the National Bank and after 1936 also the District Bank) for the period January 1921 to October 1948, inclusive. Advances were also shown throughout, both as aggregates and as a percentage of deposits, while figures of discounts and investments were added in 1923 and of Treasury deposit receipts in 1940. In February, 1949, the nine-bank series was abandoned in favour of an eleven-bank series, shown annually from 1937 and monthly from October, 1946.

The current series of bulletins, now published in the *Times Review of Industry*, show both gross and net deposits of the eleven banks annually from 1921 to the latest complete year, with monthly figures for the latest two years or so. For the latter, Lloyds Bank's seasonally adjusted index of net deposits is also shown, but on the base 1938 = 100, while for the annual figures the index has been extended back to 1921 on the same base. For the same period the main assets (cash, call money and discounts, investments and advances) are shown as percentages of gross deposits, so that a simple addition of two figures provides a series of liquidity ratios or earning assets ratios. Additional columns provide figures of notes in circulation and (from 1925 onwards) of bankers' deposits at the Bank of England, apart from more general financial information.

The *Bankers' Almanac* includes, in its Information Section, a collection of useful statistics, notably a table of changes in Bank Rate since 1844. The latter year saw the foundation of the *Bankers' Magazine*, whose files provide a convenient source of reference to various series, including figures of Scottish and Irish note circulation. *The Banker* also publishes a statistical section, useful features of which are tables showing changes in clearing bank "risk assets" and liquidity ratios, annual averages of the main clearing bank figures since 1921 and recent figures (covering several quarters) of bank advances as classified in the British Bankers' Association returns.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REVIEWS OF STATISTICAL AND ECONOMIC BOOKS

CONTENTS

	PAGE
1.—*Finney* (*D. J.*). Experimental Design and its Statistical Basis	210
2.—*Adams* (*J. K.*). Basic Statistical Concepts	211
3.—*Cohen* (*J.*) & *Hansel* (*M.*). Risk and Gambling	211
4.—*Kopal* (*Z.*). Numerical Analysis	212
5.—*Allais* (*M.*). Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine	213
6.—*Abel-Smith* (*B.*) & *Titmuss* (*R. M.*). The Cost of the National Health Service in England and Wales	214
7.—*Blacker* (*C. P.*) & *Gore* (*A. T.*). The Bethlem Royal Hospital and the Maudsley Hospital Triennial Statistical Report, 1949–51	215
8.—*Kulldorff* (*G.*). Migration Problems	216
9.—*Petersen* (*W.*). Planned Migration	216
10.—*Gilbert* (*M.*) & *Stone* (*R.*) (ed.). Income and Wealth, Series IV	217
11.—*Mehta* (*M. M.*). Structure of Indian Industries	218
12.—*Miller* (*H. P.*). Income of the American People	219

1.—*Experimental Design and its Statistical Basis.* By D. J. Finney. University of Chicago Press; London: Cambridge University Press, 1955. xi, 169 pp. $8\frac{1}{2}''$. 30*s*.

There is already a considerable number of books on the design and analysis of experiments, many of them excellent in one way or another, and so one is curious, on taking up yet another, to see what new contribution the author can possibly have to make. Dr. Finney does not disappoint us with this new addition to the series "The Scientist's Library: Biology & Medicine". On the one hand this is a book which will certainly be read with profit by those for whom it is primarily intended: namely, research workers in all branches of biological science. Many of these have not had specialized mathematical or statistical training. Nevertheless, they cannot escape quantitative studies and are in fact obliged to use some kind of experimental design, which will probably be unsatisfactory if special steps are not taken to ensure a statistically appropriate choice. On the other hand, even professional statisticians should derive some pleasure from the book, not so much for the information it contains, as for the masterly way in which it presents a wide range of modern techniques in experimental design in concise and simple language.

As Dr. Finney is at pains to point out, his object is not to provide a detailed "manual of instruction on the design and analysis of experiments but a general survey of how statistical theory can usefully guide experimental design". No specialized statistical knowledge is assumed to begin with, and so two early chapters are devoted to explaining in simple terms elementary notions such as frequency distributions, estimation, significance tests involving t and χ^2, and randomization in the simplest experimental arrangements. Then follows a more detailed treatment of randomized blocks and Latin squares. Apart from discussing general aspects of the conduct of such experiments, the comparatively simple numerical procedure needed for the analysis of variance is described in full. The decision to leave this as an indication of the sort of arithmetical analysis required in general is a wise one. Many experimenters are discouraged from learning how to choose an appropriate design, even when they are really quite capable of discriminating between relatively advanced lay-outs, because they have been unnecessarily influenced by books that present what is to them an enormously intricate mass of numerical calculations.

The remainder of the book concentrates on the design, rather than the analysis, of experiment. The general characteristics are discussed of more sophisticated designs such as balanced incomplete blocks, Youden squares, lattice designs, factorial experiments, confounding, fractional replication, split-plots and so on. It is shown that the need for such complications arises naturally in response to certain inescapable features of various experimental situations. Once the research worker appreciates how to choose the right kind of design, the main difficulty has been overcome. The question of analysis is subsidiary and may well be undertaken by an expert in this aspect of the work.

There is also a short but interesting chapter on sequential experiments, which includes not only the narrower problem of deciding when to terminate an experiment, but also the wider issue of the sequential approach to experimentation, as well as a section on experiments designed to determine optimal conditions. The next chapter contains a useful survey of the special problems

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

connected with biological assay, and the final chapter discusses various general problems involved in selecting appropriate designs, such as number of factors, choice of levels, number of replications, etc., and in properly interpreting the results of analysis.

It is greatly to be hoped that this book will be widely read by workers in the biological field, and that they will not be unduly discouraged by the price, which is rather high for the size of book, due presumably to the principal publishers being in the United States.

NORMAN T. J. BAILEY.

2.—*Basic Statistical Concepts.* By J. K. Adams. New York and London: McGraw-Hill, 1955. xvi, 304 pp. $9\frac{1}{4}''$. 41*s.* 6*d.*

The author of this book is Assistant Professor of Psychology at Bryn Mawr College. It is pleasant to observe that he has not succumbed to the temptation to write a book on "Statistics for the Psychologist", but has honestly entered the field of plain statistical texts.

The first chapter introduces "finite populations and their distributions". In place of the more usual terms "character" and "individual", "objects" are classified by "value", which is "the one and only one category into which that member is classified according to a given classification scheme", and a final definition reads; "When each member of a set has a value, the set is a *population* or *universe*". That is to say, the author considers that the term population should not be applied unless there is some way of classifying the members and, indeed, he gives as a formal definition "A *population* or *universe* is a value function". Though this is a tenable proposition, it does seem to emphasize unduly one aspect of the concept of population, and does not seem to make the development of the subject any easier.

The ideas of testing hypotheses, confidence intervals and power are developed without introducing any ideas beyond those of finite populations. Professor Neyman, in his text book on statistical inference, has expressed the view that these concepts can be introduced at an early stage in the student's training, and later experience has confirmed this view. Whether the introduction of these ideas at a still earlier stage is *generally* desirable is not certain. However it does seem likely that, for the non-mathematical class of reader the author envisages, these general concepts will be the clearer if presented free of the mathematical tangle of "infinity, continuity, density or the normal curve". The presentation of the logic of statistical inference is extremely precise and gives no indication of the perplexities in statistical theory which the ideas of power, etc., were intended to overcome, nor of the doubts and difficulties attending their introduction. The ideas of parameters and statistics are not introduced until the following chapter and this seems to be unfortunate.

The standard statistical procedures—χ^2, t and F tests and the analysis of variance are introduced in rapid succession. The author puts forward the procedure proposed by Tukey (*Biometrics*, Vol. 5, p. 99, 1949) for comparing individual means arising in analysis of variance. Although a number of alternative methods have subsequently been proposed, and in many quarters regarded as superseding Tukey's method, the latter does seem to give a reasonably useful working rule and it is encouraging to see it included in a fairly elementary text-book. There is also a good discussion on the effects of pooling sums of squares in the analysis of variance.

There is a final brief chapter on non-parametric statistics, followed by four appendices. The second of these contains more or less standard derivations of results quoted in the main text, the third and fourth give a quite useful set of statistical tables. The first is more unusual; it is on "Some hints on how to ask questions of mathematical statisticians". The main point seems to be that the problem should be "structured", i.e. put into an abstract (and presumably simplified) form. This would be very nice for the mathematical statistician but it encourages a lack of interest in applications which finally may result in too wide a gap between theory and practice. With the universal trend toward specialization this is only too likely to happen and it does not seem unduly fanciful to foresee the rise of a profession of "structurizers".

This book gives a clear outline of the simpler statistical techniques and is suitable for the embryo worker in applied fields with a strong interest in statistics. He will not gain a deep insight into the subject, or into current trends of development, but will obtain an understanding of the bases and ideas behind most common statistical procedures.

N. L. JOHNSON.

3.—*Risk and Gambling.* By J. Cohen and M. Hansel. London: Longmans, 1956. x, 153 pp. $7\frac{1}{2}''$. 14*s.*

The authors state in the preface that "this book is intended for all those who have occasion to fill in football coupons or cross a road". Its subtitle is "The Study of Subjective Probability". In it are described experiments, mostly carried out on children, designed to investigate the mental

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

processes involved in predicting the unknown, in making guesses with various amounts of prior information, etc. Studies of the subjective evaluation of vague words denoting quantity, frequency or duration, and observations of pedestrian behaviour in crossing a busy street are also described. Results are presented mainly in tabular form and deductions therefrom given in non-statistical language.

Although the book is short, the absence of a central theme makes it impossible to give a brief summary of the findings. The type of deductions made may perhaps be best illustrated by quotation. For example (p. 31): "From a variety of experiments . . . we are able to trace some of the rules of subjective probability throughout the years of development. These rules vary characteristically from year to year and differ strikingly from the rules of mathematical probability. It would hardly be an exaggeration to say that there is no such thing as a 'guess' in a sheer random sense." Again (p. 91): "So far as our evidence takes us in this experiment and within the age-range studied, we have no reason to suppose that the additive character of subjective probability differs from that of mathematical probability." Also (p. 133): "We venture the hypothesis that the stages we have distinguished in the development of the idea of a distribution partly represent mental 'structures' . . . that emerge at different ages and partly represent the fruits of experience. These 'structures' undergo changes during development. Shaped at the start by purely subjective preferences of one kind or another, they become less and less controlled by such subjective influences and assume a form which is more in accord with the objective situation."

The book should be of interest to psychologists in that new ground is broken and types of experiment capable of further development and application are described. How far the conclusions reached, sometimes based on small numbers, are valid it is difficult to assess in view of the deliberate exclusion of statistical argument. Certain of the numerical tables are hard for a layman to understand and at least one of them, Table 23, is wrongly captioned. As regards practical advice to those who fill in football coupons, the authors suggest (p. 143): "The student of subjective probability should seek those pools in which the experts themselves invest and then carefully avoid following the experts' advice." This is a rather curious mixture of subjective probability and objective fact. If subjectively one feels that the experts are likely to be right but goes counter to their advice, one is sacrificing subjective probability for the fact that the winnings will be bigger if the unexpected happens and the experts are wrong. This might be taken as a cynical commentary on the practice of filling in football coupons at all!

E. G. CHAMBERS.

4.—*Numerical Analysis*. By Zdeněk Kopal. London, Chapman & Hall, 1955. xiv, 556 pp. $8\frac{3}{4}''$. 63*s*.

Five years ago any lecturer on numerical analysis was hard pressed to recommend any books on this subject, in English, to his students. To-day, he is faced with exactly the reverse problem; the number of books published in this period on the subject makes the choice of a single book difficult.

Now Professor Kopal has added another to the list; in doing so he has not increased the difficulty of choice, but has eased it. For his book differs from most others in four important aspects.

First, the author has very wisely decided to deal with one branch of the subject exhaustively to the exclusion of all others, rather than treat all the branches in outline. The book deals with the numerical treatment of the processes of the differential and integral calculus of one real variable.

Second, the author, being an astronomer, realizes that the subject has its origins in the work of Newton, Gauss and the other giants of the past who devoted most of their working lives to astronomical calculations and therefore to a study of the methods of calculation. Many modern books, written by men recently introduced to numerical analysis through an interest in computing machines or through the ever-increasing failure of classical analytic methods in the real problems of applied mathematics, imply, if they do not actually say so, that this subject is new and arises from the invention of the automatic computing machine. This view fails to realize that the need for the calculations was the mother of the invention and this need existed 100 years ago when Babbage first attempted to produce a machine of this kind.

Third, the author has managed to write his account in a way which arouses in the reader the same thrill of discovery that the inventors of the different techniques must have felt.

Fourth, and related to the third, each topic in the book is developed by a variety of different approaches. Often these approaches lead to the same solution, but when they do not the reasons for the differences have already been disclosed in the preceding development. The inter-relationship between the various topics is also brought out by this treatment.

The central theme of the book is that the numerical treatment of analytical functions must hinge, to quote the author's words, "on the deferred approach to the limit" and on the replacement of all

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

analytic functions by polynomial approximations which form a closed system under the operations of addition, subtraction, multiplication, differentiation and integration.

Interpolation is treated as a problem of curve fitting and the various formulae are all developed from the Lagrangian form. In addition, Hermite's form in which values of the derivative as well as the function are used to define the approximating polynomial are discussed. The practical aspects (of rounding error, throw back of differences, etc.) are very capably treated.

Numerical differentiation follows easily from this and the limits of accuracy obtainable are derived for some important cases. This leads naturally to the problem of differential equations with one-point boundary conditions.

In Chapters V and VI two-point boundary condition problems are discussed. Chapter V deals with these by an algebraic approach. The solution is required at a set of points within an interval—we search for a polynomial satisfying the equation and boundary conditions at these points. This leads to a characteristic equation in the form of a determinantal equation. Chapter VI reaches this point through the variational and iterative procedures as typified by Raleigh, Ritz and Schartz respectively.

Chapter VII deals with the general problem of quadrature and starts with Gaussian quadrature and the generalization with weight functions. These formulae are progressively specialized to deal with all the more familiar Newton-Cotes formulae and so on. This reversal of the normal expositional order brings out the real reasons for many of the anomalies of these latter formulae. For example, the apparently accidentally higher precision of Simpson's rule that a mere count of ordinates would lead one to expect is really due to the central ordinate in that rule falling at a point used in a Gaussian quadrature.

This apparent digression of Chapter VII is then used in Chapter VIII to consider integral and integral-differential equations where their relationship to many-point boundary value problems is fully discussed.

While the presentation of the material is excellent, the choice of notation is rather unfortunate in many places and the enumeration of the equations by a three-level system is a great discouragement to reference back. But these are minor faults in a book which is a major contribution to the subject of numerical analysis.

K. D. TOCHER.

5.—*Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine.* By M. Allais. [Paris], Imprimerie Nationale [1955]. 55 pp. $10\frac{1}{2}''$.

In 1730 (published 1738*), Daniel Bernoulli argued that a rational man behaves so as to maximize the mathematical expectation of something, personal to him, that is now called "utility". He acknowledged the priority of Gabriel Cramer, so we may call the principle the "Bernoulli-Cramer principle of rational behaviour". In 1926, F. P. Ramsey produced rather convincing arguments in favour of the principle, and his arguments were extended and clarified by von Neumann and Morgenstern (1944 and 1947, especially 1947, pp. 617–32) and by Savage (1954). Allais is an opponent of the principle and in the thesis under review he gives his reasons and also constructs a theory that assumes less and from which less can be deduced.

Among the adherents of the principle are Marschak, de Finetti, Milton Friedman, Samuelson and the reviewer. Samuelson previously had views similar to those of Allais, but he has defected to the Bernoullians. It is a little misleading to refer to adherents of the principle as the "American school".

Allais believes that there are four considerations that a theory of rational behaviour or of risk must entertain if it is to be realistic, even as a first approximation. These considerations are (i) the distinction between monetary value and psychological value or utility, (ii) the "deformation" of objective probabilities into subjective ones, (iii) the expectation of the utility, (iv) the whole form of the probability distribution of the utility and not just the expectation. Adherents of the Bernoulli-Cramer principle regard the fourth consideration as irrelevant. The reviewer's opinion is that the only weakness of the Bernoulli-Cramer principle is that it does not take inequalities seriously enough. The principle is often insufficient to enable us to make a decision because our judgments of utilities and probabilities are usually expressed in the form of inequalities instead of being precise. (If it were not for this difficulty there would be no point in the "minimax defect" principle or type II minimax principle.†) This weakness is reminiscent of consideration (iv), above, but is by no means the same. An adherent of the principle can recognize its weakness but cannot consistently regard consideration (iv) as relevant.

* Incomplete references in this review can be completed by reference to L. J. Savage, *The Foundations of Statistics* (New York and London, 1954), reviewed in this *Journal*, 1955, *ser. A*, **118**, 245–6.

† See *J.R. Stat. Soc.*, B, **14** (1952), 107–114; and **17** (1955), 194–6, 201–2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The alleged refutation of the principle is built upon some examples on pp. 37 to 43 of the thesis, and is effectively answered by Savage (1954), pp. 98 to 104.

The thesis is a reprint of pp. 257 to 332 of *Econométrie* (Centre National de la Recherche Scientifique, Paris, 1953). It is an extract from a forthcoming larger study called *Fondements de la Théorie Economique du Risque*. It is also closely related to an article in *Econometrica*, **21** (1953), 503–546, reviewed by Savage in *Mathematical Reviews*, **15** (1954), 455. The last paragraph of Savage's review applies equally well to the present thesis and we quote it here.

"To come to the bone of the contention, suppose the man is confronted with a choice between two tickets in the same lottery. Ticket I entitles him to the prize A and Ticket II to the prize B if a white ball is drawn. Either ticket entitles him to the prize C if a white ball is not drawn. The prizes A, B, and C are themselves tickets in another lottery such that, faced with a choice about that lottery, the man would prefer A to B. Now can this man 'rationally' prefer Ticket II to Ticket I in the first-mentioned lottery? The utility theory implies that he cannot. Modern adherents of the utility theory typically either accept this conclusion as appealing, if not self-evident, or they see it as implied by other conclusions that they do accept as appealing. The author, and most others who reject the utility theory, reject the conclusion that Ticket II cannot reasonably be preferred to Ticket I and regard this conclusion as the central defect of the utility theory."

The author claims (p. 38) that prudent people, commonly considered rational, would rather take 40 francs than a half chance of 100 francs. (Could it be that with such small sums the pleasure of getting something for nothing matters more than the precise amount received? Or that prudent people often regard unnecessary gambles as sinful or as bad for their reputations? If so this example collapses as an argument against the Bernoulli-Cramer theory.)

Allais claims to have experimental evidence to show that the Bernoulli-Cramer principle is unacceptable to many sensible people. It should be held in mind, however, that optical illusions, mathematical paradoxes (e.g. Ball and Coxeter, *Mathematical Recreations and Essays*, Chapter 3) and mistakes in arithmetic can cause sensible people to contradict, in effect, mathematical theories to which they would normally give their support. The reviewer would be glad of the opportunity to perform further experiments, with the same subjects as those used by the author, but this time with real money.

It is clear that the reviewer's faith in the Bernoulli-Cramer principle of rational behaviour is unshaken by Allais's thesis. But this thesis may still be important for economics. The economist may give advice to a Government as to how to behave rationally, but he should take into account that real people do not always behave rationally. This point is familiar to modern economists (see, for example, *Uncertainty and Business Decisions* (edited by Carter, Meredith and Shackle), Liverpool University Press, 1954).

I. J. Good.

6.—*The Cost of the National Health Service in England and Wales.* By Brian Abel-Smith and Richard M. Titmuss. The National Institute of Economic and Social Research Occasional Papers. Cambridge University Press, 1956. xx, 176 pp. 10″. 27*s.* 6*d.*

In this age of the Welfare State the public services of Great Britain are so large in scope as to provide major industries in themselves. Few of us have not got friends or acquaintances who are engaged in these services. All of us are becoming increasingly conscious of having to foot the bill either by direct or indirect taxation (including contributions to National Insurance), i.e. increasingly conscious that it is not a mystical and remote State which provides the Welfare but everybody. However, many of us are not at all clear what "footing the bill" really implies in physical terms, i.e. in terms of the fiscal system used to ensure that of the goods and services available in any one year (the total social product) sufficient are diverted from the personal consumption of those who contribute to the social product to those who may not currently contribute but, by the previous creation of assets or simply by need, have created a claim upon the social product. A "ceiling on costs" in plain words is a restriction on this diversion and so if this is to be accomplished in a reasonable manner consideration has to be given to probable trends in the volume of the claim of the "dependents", and account has to be taken of the fact that contemporary changes in monetary values renders it necessary to translate costs in pounds sterling to costs in terms of actual goods and services. This kind of social accounting was the real task which the Minister of Health set the Guillebaud Committee which he appointed in 1953 "to review the present and prospective cost; to suggest means, whether by modifications in organization or otherwise, of ensuring the most effective control and efficient use of such Exchequer funds as may be made available; to advise how, in view of the burdens on the Exchequer, a rising charge upon it can be avoided, while providing for the maintenance of an adequate service; and to make recommendations". The task in so far as the accounting was concerned was undertaken by the National

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Institute of Economic and Social Research and delegated to the authors; their memorandum which was presented to the Committee forms the substance of this volume.

This task was a new one; there was no previous exercise to emulate. It was necessary to crystallize the concepts of the kind referred to above and to formulate definitions so that within the rather confused and imperfect information resources available, units could be identified which could be put together to enable the assessment of true cost to be made. How this was accomplished—the separation of capital from current expenditure, the tracing of expenditure which constituted mere book-keeping exchanges, the distinction between changes in price and changes in consumption of resources, treatment of trading activities, the breakdown in component services—hospital, executive council and local authority services—forms an invaluable methodological record. Especially is this true of the detailed analyses which appear as Appendices to the main text.

The "new look" at expenditure puts an entirely different complexion on a situation that had become highly coloured by parliamentary polemics based upon superficial book-keeping values—the current cost expressed as a share of the total national resources fell from a peak of 3·75 per cent. in 1949/50 (the first *full* year) to 3·24 per cent. in 1953/54. Contrary to public opinion the net diversion of resources to the National Health Services since 1949/50 has been of relatively insignificant proportions. The cost per head at constant prices was almost exactly the same in 1953/54 as in 1949/50. During the period under review net current expenditure on the hospital services rose while that on local authority services and executive council services fell. These trends are analysed in considerable detail.

In assessing future trends account has been taken of demographic factors, especially the increasing proportion of elderly persons (the absence of representative studies of the work of general practitioners was, here, a great handicap) and of social factors (family relationships, class, income, etc.) in determining the pattern of demand. The conclusion is drawn that "other things being equal the demand for hospital care is more influenced by the social situation of a person than by the age of a person". The authors have made a close and interesting study of the characteristics of hospital patients as a population in themselves, and of such important associated matters as the relationship of patient/staff ratios to hospital space.

The authors faced great difficulties in the lack of statistical information and they make a powerful plea "for a further step forward in the collection, classification and analysis of facts". But, and this is important, they urge that "what is now needed is refinement, analysis and operational study rather than more routine statistics". " . . . administration at most levels—regionally, locally, and at the centre—is capable of some improvement if aided with the results of a discriminating statistical service."

By their ingenuity in attempting to fill the statistical gaps Mr. Abel-Smith and Professor Titmuss have brought their existence into relief and have shown clearly what is most needed. But much more than this, they have inaugurated a new phase in the analysis of social administration and its relationship to economic development.

B. BENJAMIN.

7.—*The Bethlem Royal Hospital and the Maudsley Hospital Triennial Statistical Report*, 1949–1951. By C. P. Blacker and A. T. Gore. Triennial Statistical Report 1949–1951. xii, 115 pp. 8½″. 10*s.* 6*d.*

The amalgamation of the Bethlem Royal Hospital and the Maudsley Hospital as a Teaching Hospital for mental illnesses in July, 1948, necessitated the introduction of a revised and standardized record system suitable for both hospitals. The functions of the hospitals differed in that Maudsley had an active outpatient department and a large children's unit whereas Bethlem had neither of these services. In addition, the Joint Hospital (the contraction used in the Report) is a postgraduate teaching centre and is closely associated with the Institute of Psychiatry of the University of London. After considering the characteristics of the population served by the hospital and the inherent difficulties in obtaining a concise yet satisfactory diagnostic list, the report goes on to describe the demographic, social and economic features of the hospital adult population and the outcome of medical treatment. The children are dealt with in a separate chapter.

The total number of individual adult patients attending the hospital in the triennium was 7,787 (in-patients 2,636, out-patients 5,151), while the children numbered 1,211. Total discharges (which included patients attending for more than one period) were 8,725 adults (in-patients 3,245, out-patients 5,480) and 1,269 children (in-patients 284, out-patients 985).

The socio-economic analysis for the adults is based upon the records of the 7,787 (3,857 males and 3,930 females) new patients. There was a slight preponderance of females in both the in-patients and out-patients; in the former it was probably attributable to the availability of beds and in the latter to external causes such as the practice of referring agencies.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Males averaged 37 years and females 40 years. Out-patients were younger than in-patients. Of the male out-patients, 77 per cent. were under 45 years of age in contrast to 66 per cent. of the in-patients; for females, the respective percentages were 70·0 and 63·3.

As compared with the population of Greater London, the proportion of single persons was less among hospital patients and certain reasons for this phenomena are suggested. There was an excess of widowed males between 16 and 54 years and widowed females between 25 and 44 years in the hospital population. Divorced males between 25 and 34 years were three times, and in the age group 35–44 twice the expected number and the differences were statistically significant. For males in other ages and for all females, the differences were insignificant.

The average age at marriage of the hospital patients did not diverge greatly from that of the population of England and Wales. Similarly, the size of the family from which the patients were drawn was not dissimilar from the national experience. The number of children born to the younger patients (25–34 years) was higher than for Great Britain, which may have some clinical significance, but in other ages the fertility of the hospital sample appeared to be below the national average.

An analysis by social class indicated that among males, there were relatively more in-patients in Social Classes I and II than in Greater London, while in out-patients there was a deficiency in Classes I and II and an excess in IV and V. There were relatively few females in the highest and lowest social classes and a marked concentration in the middle classes in both categories of patient.

Information is also given about the patients' incomes, but this is limited by the incompleteness of the data; about half of the patients were in receipt of a weekly income between £4 and £7. Housekeeping claimed 40·9 per cent. of the female patients. In both sexes clerical occupations were commonest among those gainfully employed, followed, in men, by engineering and transport and communications, and, in females, by personal service. The occupations of the hospital patients corresponded with those of the population of Greater London.

Much useful information is given about admissions, the source of reference, duration of stay, special clinical investigations, diagnosis, treatment and ultimate disposal.

In contrast to the adult experience, male children outnumbered female (747 males, 464 females) and in-patients formed a smaller fraction of the total (22 per cent. for children against 37 per cent. in adults). Over half the children were 10 years of age or over, while 9 per cent. were under 5. Similar social, economic and treatment details are given for children as for adults but no diagnostic analysis is provided because a satisfactory classification of psychiatric ailments of children does not yet exist.

This is a useful and informative report and is stated to be the first of a series of triennial publications. Such a series will do much to fill a gap in knowledge of the demography of hospital patients.

A. BARR.

8.—*Migration Probabilities.* By G. Kulldorff. (Lund Studies in Geography, Ser. B., No. 14.) Lund: Gleerup, 1955. 45 pp. 9¼″. *Kr.* 5.

This is an attempt to get over differences in the size of communities when comparing rates of migration. By "migration probability" is meant the probability that a movement starting inside the region will cross the border and end outside the region. An ingenious mathematical model is set up, with parameters characterizing the size of the community and the distribution of the length of movements. The technique is tried out on the data for the community of Ashby in Sweden in the period 1935–39, and the result is not too discouraging. The author fully realizes that more extensive studies will need to be made before any general conclusions can be drawn. This study indicates that there is wide scope for the mathematical analysis of internal migration.

BRINLEY THOMAS.

9.—*Planned Migration: The Social Determinants of the Dutch-Canadian Movement.* By William Petersen. Berkeley and Los Angeles: University of California Press, 1955. London: Cambridge University Press. x, 273 pp. 9¼″. $3.50.

Professor Petersen has produced a valuable study of the recent migration from Holland, mainly from the social and demographic point of view. The background information given about Holland is particularly interesting. In spite of a high degree of urbanization according to the usual criteria, Holland has an exceptionally high rate of fertility; during this century her population has doubled, whereas that of the rest of Northern Europe has increased by less than a quarter. The birth rate in 1952 was 22 per 1,000 as compared with 16·5 for North-western Europe as a whole. Meanwhile the country has experienced a big decline in its death rate, the figure for 1952 being 7·3 per 1000.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Much attention has been given by sociologists to this problem of high fertility. It seems to arise from what Professor Kingsley Davis, in the Preface to this volume, calls "a curious demographic rivalry between Catholics and Protestants". Professor Petersen shows that " . . . the principle that children are educated in tax-supported religious schools has been established too firmly even to challenge; and this socialization has tended to perpetuate the division along religious lines and thus the authority of institutions and values that support the large-family system" (p. 40).

After a brief discussion of the sanctioned methods of relieving population pressure in post-war Holland and the conditions which have governed the recent economic upsurge in Canada, the author gives a critical analysis of migration and population growth in the two countries. He concludes that, given the factors maintaining high fertility, emigration will have little if any effect on population pressure; it is a weak substitute for birth control. The literature on the alleged "displacement" of Canadian-born population by immigrants is reviewed, and the conclusion is reached that " . . . the 20th century migration of Canadians to the United States has been an international variant of the urbanization going on within both countries" (p. 208).

The author finds much to criticize in the administrative controls set up in Holland and Canada to deal with migration, and he is sceptical about the use of economic theory as a guide to policy. It is in the social field, however, that this book makes a contribution. There are two short notes on the migration statistics of Holland and Canada and a useful bibliography.

BRINLEY THOMAS.

10.—*Income and Wealth, series IV.* Edited by Milton Gilbert and Richard Stone. International Association for Research in Income and Wealth. Cambridge: Bowes and Bowes, 1955. xii, 388 pp. 8½″ 42*s*.

This volume presents ten of the papers read to the International Association for Research in Income and Wealth at the Castelgandolfo conference in 1953. The presentation of the book maintains the high standard of this series. The book contains much of interest, but on the whole one is left with the feeling that it is a mistake to give such durable and expensive form, requiring such long preparation, to conference papers, some of which present only tentative conclusions put forward for discussion while some others are concerned with the problems of work still in progress, of which full reports appear later.

Four papers discuss the use of social accounts in economic analysis. Schouten and Lips develop an interesting accounting frame and a model for assisting government policy decisions, instruments and objectives of policy being clearly identified. One notices here and in other papers the trend towards emphasising the role of financial transactions in decision making and the space given to these in the accounts. Stuvel prefers a less mathematical approach. But it is rather sweeping to quote the lack of series free from the effects of "extraordinary phenomena, (such as the . . . Korean War)" as an obstacle to model building; these disturbances are often welcome, giving rise to greater variation and so helping in establishing relationships. His paper has many sensible things to say, particularly on productivity, the use of semi-articulated accounts, and the treatment of the terms of trade in real income measurement.

Stone's article is an annotated bibliography of the applications of social accounts in model building. Starting with national budgeting models the discussion goes on to examine the prediction of the elements of final demand; more complex models, such as Tinbergen's, are also discussed; and there are sections on input-output models and activity analysis.

The article on decision models by Frisch is a rather disconnected set of jottings, many points remaining obscure for lack of context. Starting with highly theoretical considerations about pressure coefficients and demand elasticities, he ends by presenting the frame of an actual "interflow matrix", but there is next to nothing to show the impact of the one on the other. The substance of one table (Table III) could have been given in the text in a third of the space, and section VIII seems to present little beyond the usual closed Leontieff model.

There are two papers on international comparisons of real product, one by Gilbert and Kravis on the final expenditure approach and one on the industry of origin approach by Paige and Adler. The papers partly duplicate what has been or, presumably, will be published by O.E.E.C., but they provide an opportunity for discussing some of the finer statistical problems. The paper by Gilbert and Kravis has an interesting evaluation of the methods of the final expenditure estimates, including an analysis of the product classes into three grades of reliability: very good, good and fair. Grading results is a welcome practice, and one recognizes that the judgement must be largely qualitative, particularly in this kind of work. Nevertheless it would be good if the grades were given some rough numerical standards. This would help the reader (here "very good" might mean anything from being surprised at an error of 5 per cent. to expecting errors of up to 20 per cent.), and it would also help the compiler of the estimates to crystallize his ideas when making the assessments.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

It seems that the bounds of likely error are here about 15, 30 and 60 per cent. for the three classes. This is fair enough in work on international comparisons; but the effect of optimism on grading is shown by hotels and restaurants being ranked as "very good".

Nicholson contributes a provocative article on real national income, getting back to first principles. Good points are made about the distinction of welfare and productivity measures. However, between the consumption measure at market prices and the productivity measure for individual producers at factor cost another productivity measure is introduced which excludes import duties and indirect taxes on final goods but includes indirect taxes on intermediate goods. This is supposed to measure the productivity of the economy as a whole, but the case made for it is not convincing.

Two articles, by Marczewski and Jackson, examine social accounting in Russia and in Eastern Europe.

The final paper, by Goldsmith, starts with a discussion of the problems of drawing up a national balance sheet, and goes on to present three sets of balance sheet data for the United States. First there are detailed balances for 11 sectors of the economy for the end of 1949; secondly there are series for national (combined) assets and national (consolidated) wealth for selected years since 1850; and thirdly there are series showing the growth of assets and net worth by sectors for selected years since 1900. The time series are given at current and at constant prices.

A. A. ADAMS.

11.—*Structure of Indian Industries*. By M. M. Mehta. Bombay: Popular Book Depot, 1955. xxxii, 340 pp. 9½". Rupees 22/8.

This ambitious series of studies by the Director of Economics and Statistics of the Government of Madhya Pradesh deals with certain aspects of the "structure" (a word that economists cannot resist) of Indian industry. Special attention is given to seven major industries: cotton, jute, sugar, iron and steel, coal, paper, and cement. The aspects studied are those that can be examined with the help of published company accounts and the kind of information usually given in a census of production. The result is a work replete with statistical material—much of it collected by the author and presented here for the first time—which is brought to bear on questions of fundamental interest to all economists.

The studies are divided into three parts and concern (a) the sizes of firms, (b) their location, and (c) their control and integration through multiple directorships.

The first part considers the changes in the size distribution of industrial units in recent years (a unit is defined as a group of establishments under single ownership). For the cotton industry, the comparisons made are particularly valuable in view of the length of period covered by the statistics: size distributions are given for six years in the period 1905–51. The distributions are, as usual, positively skew. The average size increases over the period, but the dispersion hardly appears to change nor to vary much among industries; the semi-interquartile range divided by the median is generally in the region of 0·2–0·4.

If firms of widely varying sizes coexist in the same industry it is natural to ask whether they are all equally efficient. The author therefore examines the "rate of return on shareholders' equity" for firms of various sizes for the decade 1938–47. He finds that the "rate of return" increases with size; for example, in the jute-mill industry the average return is 4·6 per cent. for small firms (less than 250 looms), and rises to an average of 13·2 per cent. for large firms (over 2,000 looms). This relationship is found in all the industries examined and seems to provide evidence of a general underlying tendency, despite the many qualifications that—as the author takes pains to point out—attach to the figures.

But the argument is not entirely convincing: the "rate of return" is, as far as the reviewer can understand the author's procedure, taken as the rate of dividends paid per nominal unit of share capital. This will be affected by capital reconstructions as the firm grows or declines in size, and a knowledge of their frequency and how they have been treated is essential if one is to judge the satisfactoriness of this measure of profitability. This matter is not adequately discussed in the text: but an inspection of the tables reveals that the larger and more profitable firms tend also to be the older firms. To take the jute industry again, only one of the twelve smallest firms, but eight of the nine largest, were established in the last century. So that the higher rate of return on larger firms may merely reflect a capitalization that is out of date, either because of the growth of the firm or because of the decline in the value of money.

The investigation reported in Chapter V approaches the issue from another direction. It is based on an analysis of the operating accounts of over two hundred companies according to their size. The cost components (materials, wages, etc.) are expressed as a proportion of sales for each company; and again it is found that the proportion of profit rises with the size of the firm.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The rise is by no means negligible: from a profit of about 2 per cent. for "small" companies, to about 10 per cent. for "large" companies. The difficulty in this investigation is to discover what has happened to depreciation. If it has been included together with "other expenses", the result is valid subject to the, perhaps obvious, qualification that a larger profit may be due to monopolistic advantages rather than to productive efficiency. But if, as the reviewer suspects, depreciation has been omitted from the costs of production, then a higher capital intensity among larger firms could equally account for the author's findings.

The second part of this work deals with trends in the localization of industry and follows the methods of Professor Sargant Florence in a study first published in this *Journal* in 1944. India is here divided into nine parts, and the regional distribution of workers in each of the seven industries examined is compared with the regional distribution of all workers. The years chosen for analysis are 1925, 1935 and 1945. The changes in localization in this period naturally vary from industry to industry; iron and steel, for example, hardly changes, whereas the paper industry is considerably less localized at the end of the period.

Among the many interesting results given in the third part of the book, perhaps the most forceful observation is "that nine leading families of India hold nearly 600 directorships or partnerships in Indian industries" (pp. 260–1). This section contains over thirty tables setting out various aspects of the concentration of control and the managerial links between industries. Foreign ownership and control is a matter of obvious interest in India and tabulations by community are therefore included; one of the tables shows that the proportion of British companies was 83 per cent. in 1911, 82 per cent. in 1931, but in 1951 had fallen to 62 per cent.

Many other matters are raised which the reader must explore for himself. In doing so, he will find a few grains of salt invaluable, and he would be advised not to read too much at a time. The literary style is prolix and often repetitive (pp. 79–80 and 197 are almost identical), the statistical tools are occasionally loosely handled (the frequency polygons of part I are often based on unstandardized class-intervals), and the references to other work are not always reliable. Further, the reader often misses a feeling of perspective. In part I, for example, he would like to be told—however briefly—the quantitative importance of the industries studied and of the firms included in the sample; and in part III some comparative figures of integration in other countries would have been helpful. But these are in the nature of minor blemishes in a work that will attract much attention both in India and in other countries where similar studies are being carried out.

S. J. Prais.

12.—*Income of the American People.* By Herman P. Miller. New York: John Wiley, 1955 (London: Chapman and Hall). xvi, 206 pp. 9″. 44*s*.

The American passion for collecting statistics can scarcely be better exemplified than by the figures given in this volume. It will come as a surprise to most statisticians in this country to discover that the American Population Censuses of 1940 and 1950 included questions about such a "touchy" subject as personal income. In the 1940 Census all persons of 14 and over were asked to give the amount of money wage or salary received by them during 1939 and to say whether or not they had received $50 or more from all other sources combined. In the 1950 Census a sample of one in five persons of 14 and over were asked to give the amounts of money income received during 1949 under three headings: (1) wage or salary, (2) income from self-employment, (3) other sources, including interest, dividends, rent, pensions and allowances. Moreover, if the person concerned was the head of a family he was also asked about the incomes (from all sources) of the other members of his family. The one limitation was that people who received more than $10,000 from any source were not obliged to report the exact amount received.

The mass of data on personal incomes which was assembled in this fashion can be used for very elaborate statistical analysis. Dr. Miller's book, which is the first study to make use of the 1950 Census results, is a useful contribution to a deeper understanding of the factors influencing the distribution of income. Some of his observations, it is true, are not very original and merely repeat, or quote from, the conclusions of earlier investigators. The most interesting, and original, part of his work consists of an analysis of the distribution of income within occupations. Since he can draw upon a complete census or, at the worst, a 3⅓ per cent. sample of the population, he can break his data down into unusually homogeneous groups. In his appendix tables he publishes income frequency distributions and several parameters, including means and medians, for a long list of occupational groups, males and females being shown separately.

Dr. Miller has tried to use this new information to throw some light on the underlying causes of the shape of the size distribution of income. But this is treacherous ground in which the search for mathematical patterns has often taken the place of economic reasoning. Dr. Miller is able

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to show that when the aggregate distribution curve is broken down into its component curves—for each occupational group—these component curves are more symmetrical than the aggregate curve. Dr. Miller is inclined to exaggerate the degree of symmetry obtained in the individual curves and to suggest that, if one could have entirely homogeneous occupational groups the curves would approximate to normal curves. There is, of course, nothing essentially strange in the discovery that the more homogeneous the sub-groups the smaller the relative dispersion and the greater the degree of symmetry. In an almost perfect market, where all men of equal ability could command a very nearly equal wage, one would expect the distribution of wages for narrowly defined occupational groups to be little dispersed and, perhaps, to be symmetrical. In other words, the forces of supply and demand would be pushing the wage of each man in a given occupation towards his equilibrium wage, and such small deviations as remained would be due solely to temporary and accidental imperfections in the market, whose effects might well be equally distributed on either side of the mean. In practice, however, even with Dr. Miller's huge sample the occupational groups cannot be narrowly enough defined to test this hypothesis. The distributions which he publishes for manual workers' wages are reasonably symmetrical, but they do not approximate closely to the normal form. In general, they approach a normal distribution for frequencies below the mode and a log-normal distribution for those above the mode. Any satisfactory hypothesis to explain the distribution of income must, in my opinion, treat these two limbs of the distribution separately.

Dr. Miller discusses many other aspects of income distribution in the United States and he produces interesting figures comparing the incomes of town and country workers, negroes and whites, young and old, and so on. In this part of the book his conclusions serve, on the whole, rather to confirm earlier investigations than to break new ground. Nevertheless, it is useful to have in one compact volume so much carefully compiled information on this very complex subject.

HAROLD LYDALL.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL NOTES

(1) British Official Statistics

The interim index of retail prices published by the Ministry of Labour and National Service which was 154 in December, 1955 (prices at June 17th, 1947 = 100) fell to 153 in January when it was calculated for the last time on the old method. A new index started with the base date January 17th, 1956, which measures changes in the average level of prices of the commodities and services purchased by the great majority of households, i.e. by practically all wage-earners and most small and medium salary earners. The number of items priced is about 350, over 80 more than the number used for the previous index. The weighting system has been computed from information provided by the household expenditure inquiry made in 1953–4 adjusted to correspond with the level of prices ruling in January 1956. As a supplementary calculation the Ministry will continue to publish a figure on the basis of June, 1947 = 100, by linking together the old and the new series. At February 14th, 1956, the new index stood at 100 and the supplementary figure was 153, both showing no change compared with the preceding month. At March 13th, 1956, the figures were 101·3 and 155. The rise in the index in March was due mainly to higher prices for bread, eggs, potatoes, fruit and bacon.

The detailed figures of the weights used in calculating the new index and the indices for the different commodity groups at February 14th and March 13th are as follows:

	Food	*Alcoholic Drinks*	*Tobacco*	*Housing*	*Fuel and Light*	*Durable Household Goods*	*Cothing and Footwear*	*Transport and Vehicles*	*Miscellaneous goods*	*Services*	*All Items*
Weights:	360	71	80	87	55	66	106	68	59	58	1,000
Indices at											
Feb. 14th, 1956	99·6	100·0	100·0	100·4	100·0	100·1	100·2	100·1	101·1	100·1	100·0
Mar. 13th, 1956	102·9	100·0	100·0	100·5	100·0	100·3	100·4	100·3	101·1	101·4	101·3

The new index is based on proposals made in a Report of the Cost of Living Advisory Committee. The Committee state that they have examined for the purpose the information already available relating to the results of the enquiry into household expenditure made in 1953–4. Subsequently a full report on this enquiry will be published.

The Ministry of Labour index of weekly wage rates, which was 154 in December (June 30th, 1947 = 100) rose to 156 in January, to 157 in February, and to 161 in March. The principal classes of workers whose rates of wages were increased during the period, included building and civil engineering workers, agricultural workers, railwaymen, engineers, shipbuilders, coal miners, and workers in the electricity supply industry and chemicals.

The six-monthly enquiry into average earnings showed that in October, 1955, the average earnings per week in the industries covered by the enquiries were 222*s.* 11*d.* for men, 115*s.* 5*d.* for women, 94*s.* 2*d.* for youths and boys and 75*s.* 8*d.* for girls. These figures represent increases over October, 1938 of 223, 255, 261 and 309 per cent. For all workers combined the average was 187*s.* 2*d.*, an increase of 251 per cent. The index for October, 1955, compared with 100 at April, 1947, was 181. This compares with the weekly wage *rate* index of 148.

The average hours worked per week in October, 1955, were 46·9 (48·9 for men only). In manufacturing industries alone the average earnings in October, 1955, were 231*s.* 1*d.* for men and 116*s.* for women. The corresponding hourly earnings were 56·9*d.* and 33·3*d.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The total working population and the numbers in civil employment in the four months ended February, 1956, were as follows:

(*Thousands*)

Date	*Total Working Population*			*Numbers in Civil Employment*		
	Males	*Females*	*Total*	*Males*	*Females*	*Total*
Nov., 1955 .	16,132	7,973	24,105	15,225	7,881	23,106
Dec., 1955 .	16,122	7,896	24,018	15,199	7,799	22,998
Jan., 1956 .	16,141	7,879	24,020	15,208	7,780	22,988
Feb., 1956 .	16,130	7,860	23,090	15,192	7,761	22,953

The number of persons on the unemployment registers of the Employment Exchanges rose by 48,000 in January and by 12,000 in February. In March the figure fell by 10,000. The total for March represented 1·2 per cent. of the estimated total of employees in Great Britain. The percentage in the different Regions ranged from 0·7 in the Midlands and North Midlands to 1·7 in the Northern Counties, 2·6 in Scotland and 2·0 in Wales.

The following is a sex analysis of the figures:—

Number of Unemployed Persons on the Registers of Employment Exchanges

Date	*Men and Boys*	*Woman and Girls*	*Total*
Dec. 6th, 1955 . .	139,006	76,644	215,650
Jan. 16th, 1956 . .	173,176	90,734	263,910
Feb. 13th, 1956 . .	186,455	89,173	275,628
Mar. 12th, 1956 . .	175,903	89,619	265,522

Of the total of 265,522 in March, 72,323 had been unemployed for not more than 2 weeks, 63,310 for two to eight weeks, and 102,444 for over 8 weeks, while 27,445 were temporarily stopped.

The numbers of insured persons absent from work owing to sickness, including self-employed as well as employed, were 860,900 in December, 1,079,100 in January, 1,195,800 in February and 1,062,600 in March. The numbers of employed persons absent owing to industrial injuries were 55,300, 62,300, 65,800 and 61,800.

Figures published in the *Ministry of Labour Gazette* for February show that the estimated total number of employees in Great Britain at the end of May, 1955, was 21,460,000 (13,960,000 males and 7,500,000 females). These figures showed increases of 120,000 males and 150,000 females compared with a year earlier. The United Kingdom totals for May, 1955, were 14,257,000 males and 7,676,000 females.

In obtaining the information as to average earnings referred to above, the Ministry of Labour also ascertained the propoortion of workers paid under systems of payment by results. These showed that in all the industries covered the percentage for men rose from 18 in 1938 to 29 in October, 1955. In the case of women the percentage fell from 46 to 42. For all workers combined there was a rise from 25 to 32 per cent. The latter figure has been stationary since 1951. In manufacturing industries alone the percentage for men is now 38, compared with 29 in 1938, for women it is 45 compared with 48 and for all workers 40 compared with 33.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The General Register Office have published the results of the last two years of an enquiry designed to test the practicability of using general practitioners' clinical records as a source of morbidity statistics.* This is a continuation of a previous study of the first year's results which gave an outline of the objectives of the enquiry and the methods used.

The enquiry covers 13 practitioners in 10 practices, with lists totalling some 37,000 patients. The various tables analyse sickness experience by disease in relation to sex, age and separate practices according to the number of patients affected and the numbers of medical consultations. Subsidiary analyses show referrals to hospital, certificates issued, place of consultation and frequencies of consultations and illnesses. The experience gained in this enquiry has helped to pave the way to a joint survey by the College of General Practitioners and the General Register Office in which some 170 practitioners are now engaged in voluntarily keeping standard records for twelve months and making them available for statistical analysis. This new enquiry covers nearly 1 per cent. of the population of England and Wales and will provide much useful information to both the medical research worker and the administrator.

(2) Other Statistics

The United Nations *Statistical Yearbook* 1955 (New York, pp. 644) in addition to the usual features, includes for the first time an index of the prices of primary commodities in international trade for the period 1948–1955 (base 1953 = 100) distinguishing between agricultural mineral and food products, calculated by the U.N. Statistical Office. Many tables have been extended or completed. In particular the tables on total population by sex, economically active population, housing and illiteracy, due to the data now becoming available from the censuses of population taken around 1950, and statistics of the production of coal, petroleum, electrical energy, cement, pig-iron and steel for pre-war and post-war years in the U.S.S.R. have been provided for the first time by the Government. All index numbers for this issue have been, it is stated, "re-calculated" wherever possible, to the base 1953 = 100. The method of recalculation is not given, apparently it is the method, not always satisfactory, of simple division of the index numbers ($\times$ 100) by the 1953 figure.

In *The Year Book of International Trade Statistics*, 1954 (U.N. publication, U.S. $5, available from H.M. Stationery Office) the United Nations has brought together statistics relating to the trade of some one hundred countries, covering about 98 per cent. of world trade in 1954. The information given for each country falls under three main headings: (i) historical series for the years 1930 to 1954; (ii) tables showing the quantity and value in national currencyof imports and exports, in recent years, analysed by commodities; (iii) tables showing, for recent years, the value in national currency of import and export trade, analysed by principal countries of provenance and destination.

The figures are obtained from official sources published by the Governments of the countries concerned; it is, thus, an orderly collection of facts and not an attempt to interpret them. Some trends can, however, quickly be appreciated by reference to the tables dealing with world trade. It is seen that the value of world trade in 1954 was more than three times that of 1937, measured in U.S. dollars, with the exports of the highly industrialized areas of the world (United States, Western Europe and Japan) contributing considerably more than their share to the increase in quantity, and markedly less than their share to the rise in price over that period. The figures also show that a higher proportion of the exports of manufacturing areas than before the Second World War is now sent to less developed areas, the proportion having risen from 40 to 45 per cent. between 1937 and 1954.

Among the areas for which figures are given, the largest increases in exports seem to have been those of the Middle East (405 per cent. higher in 1954 than in 1937, mostly due to increased exports of oil) and North America (which shows a rise of some 337 per cent.). The smallest increase both in exports and imports is shown by the Far East (exports up 65 per cent. and imports up 123 per cent.).

World exports had a value of 75,100 million U.S. dollars in 1951, having climbed from 55,400 m. in 1950. After falling in 1952 they began to climb again and by 1954 had reached 76,250 million U.S. dollars.

---

* *Studies on Medical and Population Subjects, No.* 9—General Practitioners' Records. An analysis of the clinical records of some general practices during the period April, 1952, to March, 1954. Price 6*s* 6*d.* net (or by post from P.O. Box 569, London, S.E.1), price 6*s* 9½*d.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The figures bring into relief the remarkable growth in the exports of the United States, which now account for some 20 per cent. of world exports as compared with 14 per cent. in 1937. By contrast United States imports have grown less rapidly, amounting to some 13 per cent. of all imports in 1954 as compared with 12 per cent. in 1937. This disparity between the levels and rates of growth in the United States exports and imports helps to explain some of the current problems of international currency convertibility.

The share of Continental Western Europe, including dependencies of countries in Western Europe, in world exports, has recovered steadily year by year in the post-war period to a figure of 32 per cent. of world exports in 1954 as compared with 31 per cent. in 1937. The exports of the sterling area countries in 1954 constituted 24 per cent. of world exports, compared with 26 per cent. in 1937, a fall mainly attributable to the decrease in the exports of the United Kingdom and Malaya (both down 1 per cent. over this period).

An important advantage of a Year Book of this kind is the way in which it facilitates rapid comparisons between different countries in respect of their importation of individual and perhaps minor commodities of particular interest to individuals or firms, whether these be hand tools, travel goods or watches and clocks. For most of these commodities information is available, often in terms both of weight (in metric tons) and value (in the local currency).

The *Economic Survey of Europe in* 1955 (the ninth report of the Economic Commission for Europe, Geneva, 1956, 350 pp., 18*s*.) points out that the year 1955 was another year of rapid economic expansion which took place without any major changes in price levels. As in previous reports, the countries of Western Europe are treated separately from those of Eastern Europe (including the U.S.S.R.) but this year two special chapters are added, one devoted to the labour market problems of Western Europe examined against the background of the present high levels of employment and one to the investment problem, that is the obtaining of a level and pattern of investment which will provide for a satisfactory long-term rate of expansion in the economy. In addition to the great wealth of material contained in the statistical tables, a special feature in this year's report is the increase in the data obtained from the countries of Eastern Europe and the U.S.S.R.

The *Annual Epidemiological and Vital Statistics* Report of the World Health Organization for 1952 has been published (W.H.O., Geneva 1955; £2 10*s*.) and contains tables of vital statistics for such countries in which the medical and administrative organizations permit the preparation of reliable figures. (The series has included four previous numbers covering 1939–1946, 1947–1949, 1950 and 1951, and is a continuation of the "Annual epidemiological reports" of the Health Organisation of the League of Nations for the years 1922 to 1938). The tables give for each country populations by age and sex, births, deaths, deaths under one year and under 4 weeks, natural increase, actual numbers and rates; mortality rates by age and sex; deaths by cause, in sex and age groups; cases of and deaths from a range of notifiable diseases. The volume is accompanied by a carefully worded introduction concerning the coverage and validity of the data.

The 1954 edition of *National Income* which is published by the Office of Business Economics of the United States Department of Commerce and forms a supplement to the *Survey of Current Business* has recently been published. The present volume reviews shifts in the use of national output, changes in income flow, expansion of personal income and fluctuations in economic activity. Considerable space is devoted to the nature and significance of the income data, to the construction of national output measures and the definitions of concepts and terms. The detailed coverage is wide and includes sections on wages and salaries, contributions for social insurance, corporate profits, interest, personal consumption, new construction, net foreign investment, personal saving and capital consumption allowances. The present report contains all the national income statistics of the Office of Business Economics with minor exceptions and these supersede all previously published figures; it is proposed to keep the series up-to-date in the monthly *Survey of Current Business*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL AND ECONOMIC ARTICLES IN RECENT PERIODICALS

UNITED KINGDOM—

Accounting Research—

January 1956—The measurement of profits—a study in methods. III—the profits index: *R. K. Nigam.* Accounting dynamics—IV: *F. Sewell Bray.*

Advancement of Science—

March 1956—Internal migration in England and Wales 1951: *R. H. Osborne.*

British Journal of Preventive and Social Medicine—

January 1956—Heights and weights of Scottish schoolchildren: *J. A. Grant Keddie.* Selection of miners: a survey of school-leavers in a village in South Wales: *I. T. T. Higgins, P. D. Oldham, A. J. Merrick* and *M. I. Dunsdon.* Social medicine studies based on Civilian Medical Board records. V. Physical and occupational characteristics of men with hernia, spinal curvature, and rheumatism: *A. M. Stewart, J. W. Webb* and *D. Hewitt.*

British Journal of Sociology—

March 1956—Family allowances: *N. D. Vandyk.* Employment of sociology and anthropology graduates 1952 and 1954: *J. A.* and *O. L. Banks.*

Economic Journal—

December 1955—The teaching of economics in schools and universities: *L. C. Robbins.* Interactions of cycles and trends: *B. Higgins.* Stability and growth: *H. Brems.* Sequence of analysis and the theory of rate of interest: *A. Ll. Wright.* Some theoretical implications of international aspects of the United States recession, 1953/54: *T. Balogh.*

Economica—

February 1956—Industrial growth and industrial revolution: *D. C. Coleman.* Marshall's demand curve: *R. F. G. Alford.* Earned and investment incomes, U.K. 1952–53: *E. C. Rhodes.*

Journal of Agricultural Economics—

January 1956—Labour in Scottish agriculture: *G. F. Hendry.* Linear programming and the cost of pig fattening rations: *H. C. Mackenzie* and *T. E. Godsell.*

Manchester School of Economic and Social Studies—

January 1956—The climacteric of the 1890's: a critical note: *D. J. Coppock.* Some aspects of Building Society finance: *G. W. Murphy.* A women's clothing multiple: *M. J. Wray.* A note on managerial income in retail distribution: *S. Pollard.* The development areas: a further note: *J. H. Dunning.*

Operational Research Quarterly—

December 1955—Generalizations of the warehousing model: *A. Charnes* and *W. W. Cooper.*

Oxford Economic Papers (New Series)—

February 1956—The Raw Cotton Commission, 1948–52: *J. Wiseman* and *B. S. Yamey.* Benelux and problems of economic integration: *W. Robertson.* Growth of employment and diversification of industry in the Development Areas: *P. Robson.* Marshall and the trade cycle: *J. N. Wolfe.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Oxford University Institute of Statistics Bulletin—

February 1956—Consumers' net worth: the 1953 Savings Survey: *K. H. Straw.* Gross national product of Wales, 1950: *E. Nevin.* Has the distribution of income become more unequal?: *D. Seers.* A "human needs" diet: autumn 1955: *T. Schulz.*

Population Studies—

March 1956—Population trends and population policy in Ceylon: *N. K. Sarkar.* A study of demographic trends in West Bengal during 1901–1950: *A. Ghosh.* The demography of Chile: *O. Cabello.* A critique of the Glass-Grebenik model for indirectly estimating desired family size: *R. J. Potter, Jr.* Some remarks on coverage checks in population censuses: *S. S. Zarkovic.*

Royal Society of Edinburgh, Proceedings (*Section A*)—

Vol. LXIV, Part II—A statistical examination of the centroid method: *D. N. Lawley.* Studies in practical mathematics. VIII. On the iterative methods of Lin and Friedman for factorizing polynomials: *A. C. Aitken.*

Sociological Review—

December 1955—The family life of old people: an investigation in East London: *P. Townsend.*

AUSTRALIA AND NEW ZEALAND—

Economic Record—

November 1955—An empirical micro-study of some factors influencing farm net investments: *R. A. Pearse.* Wool prices and pastoral policies, 1867–75: *A. Barnard.*

INDIA—

University of Baroda, Journal of the Maharaja Sayajirao—

October 1955—The square triangular matrix with unit elements: *N. M. Bhatt.* Simultaneous confidence bounds for certain parameters in the analysis of variance situation: *K. V. Ramachandran.* Derivation of Sheppard's corrections from Craig's result: *C. G. Khatri.*

Sankhyā—

October 1955—On Fisher's lower bound to asymptotic variance of a consistent estimate: *G. Kallianpur* and *C. R. Rao.* Hypotheses concerning the means of observations in normal samples: *P. B. Patnaik.* On a bias in a crop-cutting experiment (application of integral geometry to areal sampling problems—Part V): *M. Masuyama* and *J. M. Sengupta.* On statistics independent of a complete sufficient statistic: *D. Basu.* On the power function of a test of significance for the difference between two proportions: *C. C. Sekar, S. P. Agarwala* and *P. N. Chakraborty.* Multidimensional hypergeometric distribution: *S. Ram.* A new test of business trends: *K. Wagner.* The production function for Indian manufactures: *M. M. Dutt.* Shortage in retail handling with special reference to handling of rice: *K. N. Chatterjee.*

UNITED STATES—

American Economic Review—

March 1956—Agriculture in the Nation's economy: *J. D. Black.* Marginal policies of "excellently managed" companies: *J. S. Earley.*

American Statistician—

December 1955—Federal statistical activities. The Gleichschaltung of the German Statistical Yearbook: *E. C. Hughes.* The early teaching of statistics in American colleges and universities: *P. J. Fitzpatrick.* Questions and answers: *E. Rubin.*

February 1956—Multiple regression and correlation analysis on the IBM type 701 and type 704 electronic data processing machines: *F. S. Beckman* and *D. A. Quarles, Jr.* Admissibility in court of Univac seasonal computations: *W. F. Ferger.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Annals of Mathematical Statistics—

March 1956—Large sample theory: parametric case: *H. Chernoff*. A "mixed model" for the analysis of variance: *H. Scheffé*. Keeping moment-like sampling computations simple: *J. W. Tukey*. Symmetric functions of a two-way array: *R. Hooke*. Some applications of bipolykays to the estimation of variance components and their moments: *R. Hooke*. On the estimation of regression coefficients of a vector-valued time series with a stationary residual: *M. Rosenblatt*. An application of information theory to multivariate analysis: *S. Kullback*. On the characteristics of the general queueing process, with applications to random walk: *J. Kiefer* and *J. Wolfowitz*. Tolerance regions: *D. A. S. Fraser* and *I. Guttman*. Generalized tolerance limits: *J. H. B. Kemperman*. On a characterization of the stable law with finite expectation: *R. G. Laha*. A variable probability distribution function: *R. E. Wheeler*. Uniform convergence of random functions with applications to statistics: *H. Rubin*.

Bell System Technical Journal—

January 1956—Statistical techniques for reducing the experiment time in reliability studies: *M. Sobel*. A class of binary signaling alphabets: *D. Slepian*. Transistor circuits for analog and digital systems: *F. H. Blecher*.

March 1956—Theories for toll traffic engineering in the U.S.A.: *R. I. Wilkinson*.

Biometrics—

December 1955—Partially replicated Latin squares: *W. J. Youden* and *J. S. Hunter*. The relative size of the inter and intra-block error in an incomplete block design: *W. A. Thompson, Jr.* A note on design and analysis of soil insecticide experiments: *P. Sprent*. Covariance analysis as an alternative to stratification in the control of gradients: *A. D. Outhwaite* and *A. Rutherford*. On the analysis of variance of a two-way classification with unequal subclass numbers: *C. Y. Kramer*. A method of analysis for a double classification arranged in a triangular table: *N. Ditchburne*. Appropriate scores in bio-assays using death-times and survivor symptoms: *J. Ipsen*. Matrices in quantal analysis: *P. J. Claringbold*.

March 1956—Fractional replication for mixed series: *M. Morrison*. Block effects in the determination of optimum conditions: *R. M. DeBaun*. Adjustment by covariance and consequent tests of significance in split-plot experiments: *J. T. Truett* and *H. Fairfield Smith*. A note on the combination of estimates of relative potency in multiple assays: *P. M. Clarke*. Missing and "mixed-up" frequencies in contingency tables: *G. S. Watson*. Contributions to simultaneous confidence interval estimation: *K. V. Ramachandran*. Random genetic drift in a tri-allelic locus: exact solution with a continuous model: *M. Kimura*. Multivariate analysis and agricultural experiments: *D. J. Finney*. The relation between quantal and graded responses to drugs: *P. S. Hewlett* and *R. L. Plackett*. One likelihood adjustment may be inadequate: *H. W. Norton*.

Econometrica—

January 1956—Long-term trends in food consumption: a multi-country study: *L. Jureen*. Solutions of saddle value problems by differential equations: *T. Kose*. Preference, choice and the theory of games: *T. Majumdar*. Dynamic programming under uncertainty with a quadratic criterion function: *H. A. Simon*. Report of Inter-Society Committee on the mathematical training of social scientists: S.S.R.C. Committee Report.

Harvard Business Review—

March–April 1956—The stock market in perspective: *J. F. Weston*. Guides to inventory policy. II. Problems of uncertainty: *J. F. Magee*.

Industrial Quality Control—

January 1956—Quality control methods for two related values: *J. E. Jackson*. Corrective action in a quality control program: *H. I. Matosoff*. Some uses of statistics in plant maintenance: *J. Armstrong, Jr.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Journal of Political Economy—

February 1956—India and China: development contrasts: *W. Malenbaum.* The problem of economic assumptions in mathematical economics: *D. Bodenhorn.* The statistics of monopoly and merger: *G. J. Stigler.*

Mathematical Tables and Other Aids to Computation—

January 1956—Some Monte Carlo experiments in computing multiple integrals: *P. Davis* and *P. Rabinowitz.* Generating and testing pseudo random numbers on the IBM type 701: *D. L. Johnson.* Evaluation of an integral arising in numerical integration near a logarithmic singularity: *Y. L. Luke.*

Milbank Memorial Fund Quarterly—

January 1956—Problems of trend determination during a transition in fertility: *N. B. Ryder.* Some aspects of fertility in Egypt: *M. A. El-Badry.* Social and psychological factors affecting fertility. XXX. Extra-familial participation of wives in relation to interest in and liking for children, fertility planning, and actual and desired family size: *L. Pratt* and *P. K. Whelpton.* The effects of changes in mortality and fertility on age composition: *A. J. Coale.*

Operations Research—

February 1956—Statistics and operations research: *P. M. Morse.* Operations research in production control—a progress report: *A. Vazsonyi.* Shock waves on the highway: *P. I. Richards.* Criteria for the selection of water-resource projects: *R. N. McKean.* The traveling-salesman problem: *M. M. Flood.* Machine repair as a priority waiting-line problem: *T. E. Phipps, Jr.*; Comments: *W. R. van Voorhis.* A polar-planimeter method for determining the probability of hitting a target: *G. R. van Brocklin, Jr.* and *P. G. Murray.* Prediction of odd-lot stock transactions: *H. K. Weiss.* Production scheduling by the transportation method of linear programming: *E. H. Bowman.* A competitive-bidding strategy: *L. Friedman.*

Psychometrika—

December 1955—*F*-test bias for experimental designs of the latin square type: *N. Gourlay.* Characteristics of two measures of profile similarity: *C. W. Harris.* The estimation of the discriminal dispersion in the method of successive intervals: *R. H. Burros.* The law of comparative judgment in the successive intervals and graphic rating scale methods: *H. J. A. Rimoldi* and *M. Hormaeche.* A statistical model for relational analysis: *J. D. Luce, J. Macy, Jr.* and *R. Tagiuri.*

Quarterly Journal of Economics—

February 1956—Social indifference curves: *P. A. Samuelson.* Monetary policy in a changing world: *E. Miller.* A contribution to the theory of economic growth: *R. Solow.* Tariffs, income and distribution: *M. C. Kemp.* Prices, wages and industrial productivity in Australia and New Zealand: *E. Lerdau* and *J. Rowe.*

Review of Economics and Statistics—

February 1956—Wages and foreign trade: *I. B. Kravis.* Distance and the pattern of intra European trade: *W. Beckerman.* The United States cycle in private fixed investment 1946–50: *C. A. Blyth.* Efficiency and pricing in the coal industry: *J. M. Henderson.* Regression methods of estimating agricultural income of counties: *J. L. Fulmer.* Equality of income distribution and consumption expenditures: *W. J. Bilkey.*

AUSTRIA—

Statistische Vierteljahresschrift—

Vol. VIII, Part 3—Reallohnvergleich Österreich-Deutschland: *J. Pfanzagl.* Moderne statistische Methoden in der Medizin: *F. Brandstetter.* Zur Ermittlung der Blitzfährdung: *G. Bruckmann.* Inwieweit sind die Erhebungen der LBG in Wien repräsentative für die gesamte Landwirtschaft Österreichs?: *R. Herzner.* Über die Berechnung der mittleren Dauer der Arbeitslosigkeit: *J. Pfanzagl.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

DENMARK—

Nationaløkonomisk Tidsskrift—

Vol. 93, Parts 5–6—The role of persuasion in economic affairs: *Sir D. Robertson.* Linear programming i produktionsteorien. II: *S. Danø.*

FRANCE—

Journal de la Société de Statistique de Paris—

July–September 1955—La situation du patrimoine national bâti entre 1850 et 1950: *M. Hermitte.* Recherche statistique de facteurs indépendants: *P. Delaporte.* Chronique des statistiques bancaires et des questions monétaires: *P. Cauboue.* Un problème de sondage parmi les éléments dont la distribution est très dissymétrique: *P. Thionet.*

October–December 1955—Introduction à l'étude des séries cumulées: *L. Amy.* Le rôle de la statistique dans la recherche opérationnelle: *G. Kreweras.* Distribution d'échantillon et de caractéristiques d'échantillon quand la population de référence est Laplace-Gaussienne de paramètres inconnus: *A. Laurent.*

Revue de Statistique Appliquée—

Vol. III, No. 3—Détermination simultanée de l'engagement réel de 450 machines: *E. Kleinmann.* Comparaison de trois réfractaires pour plaques de coulées du point de vue de leur comportement en service: *F. Gottardi* and *A. Palazzi.* L'analyse de la variance et les plans factoriels dans leurs applications à l'étude de la conservation des agrumes: *L. Pillet.* Contrôle statistique des mises au mille dans la coulée et le laminage de l'acier: *M. Spreux.* Étude des délais d'attente des Aéronefs à l'atterrissage: *G. Pilé.* Choix et détermination pratique d'intervalles de confiance: *M. Dumas.* Nombre de machines automatiques à confier à un ouvrier en vue d'atteindre le prix de revient industriel minimum: *E. Daru.* Une nouvelle méthode d'exploitation des résultats expérimentaux: *G. Blet.*

Vol. III, No. 4—Étude du phénomène de choc—cas particulier du lucoflex dans sa zone de transition: *J. Lesavre.* Application de la carte de contrôle a la surveillance du taux en filature peignée—efficacité du test—centrage initial: *F. Monfort.* Mesure et contrôle de la dispersion à partir de groupes de deux: *B. Cyffers* and *A. Vessereau.* Étude: traitement statistique du comportement des véhicules au carrefour: *B. P. Adhikari.* Décisions à propos de sondages: *P. Thionet.* Sur l'efficacité des plans statistiques d'expérimentation: *F. Bastenaire.* L'analyse des groupes à variables multiples ou analyse typologique: *A. Vidal.*

GERMANY—

Allgemeines Statistisches Archiv—

Vol. 39, Part 3—Eine Verallgemeinerung des einfachen Urnmodells und ihre Anwendung in der Stichprobentheorie: *H. Kellerer.* Stichprobenmethoden in der quantitaven Marktforschung: *D. Strobel.*

Vol. 39, Part 4—Der Beitrag der Bevölkerungsstatistik zur Untersuchung der Zusammenhange zwischen Bevölkerung und Wirtschaft: *H. Schubnell.* Die statistische Erfassung des Berufes und der sozialen Stellung für ökonomische und soziologische Untersuchungen: *F. Zopfy.* Die Haushaltung als wirtschaftende Einheit und ihre statistische Erfassung: *K. Szameitat.*

Mitteilungsblatt für Mathematische Statistik—

Vol. 8, Part 1—Überführung der Schwerpunkts—in die Hauptachsenlösung und weitere Entwicklung der Näherungsverfahren in der Faktorenanalyse: *R. Bargmann.* Grenzwertsätze f. verschwindende Übergangswahrscheinlichkeiten: *W. Kanngiessar.* Prüfverfahren f. Variable mit weitem under engem Toleranzbereich: *K. Brücker-Steinkuhl.* Über die stochastische Theorie der Merkmalsiterationen: *O. Ludwig.* Anwendung der logarithmischen Normalverteilung: *R. Wartmann.*

Weltwirtschaftliches Archiv—

Vol. 76, Part 1—Die quantitative Wirkungen einer Wiederherstallung der Währungskonvertiebilität: *W. Gatz.* The industrialization of the Lebanon and Syria: *K. Grunwald.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ITALY—

Giornale degli Economisti e Annali di Economica—

September–October 1955—Le equazioni dell'equilibrio statico economico-finanziario in taluni casi complessi: *G. Sensini.* Grandi mezzi per le grandi decisioni di politica economica: *S. Ricossa.*

November–December 1955—Su alcune caratteristiche degli sviluppi polinomiali di probabilità: *G. Demaria.*

JAPAN—

Institute of Statistical Mathematics, Annals—

Vol. VI, No. 3—Estimation of linear regression coefficients in time series: *A. Mine.* Note on a relation between the distribution functions and characteristic functions: *M. Motoo.* Note on the Neyman-Pearson's fundamental lemma: *Y. Suzuki.* Bounded significance level tests for comparing quantities of two possibly different continuous populations: *J. E. Walsh.* New aspects of von Neumann's model with special regard to computational problems: *H. Nikaidô.* On the statistical investigation of diagnosis in the internal medicine: *S. Sakino* and *U. Hirata.*

Vol. VII, No. 1—Some problems of sampling in the forest survey: *K. Matusita, C. Hayashi, M. Isida, H. Uzawa, H. Hudimoto, H. Akaike* and *T. Uematu.* On the theory of sampling with probabilities proportionate to given values: *S. Yamamoto.* The significance of the discordant variance estimates: *M. Siotani.* Note on the moments of the logarithmic non-central χ^2 and Z distributions: *B. M. Bennett.* On the joint distribution of the mean and standard variation: *B. M. Bennett.*

Union of Japanese Scientists and Engineers, Reports of Statistical Application Research—

May 1955—On the power function of Welch's test procedure in the two sample problem: *S. Ura.* On a certain queueing process: *T. Homma.* A note on formulas for finding estimates for any number of missing plots: *E. Ishikawa.*

August 1955—A normal approximation to Poisson distribution: *H. Makabe* and *H. Morimura.* A normal approximation to binomial distribution: *H. Makabe.* Notes on the confidence band of population distribution: *K. Kunisawa, H. Makabe* and *H. Morimura.* Notes on statistical applications of information theory: *M. Sakaguchi.* Elementary inequalities for Mills' ratio: *Y. Komatu.*

NETHERLANDS—

Statistica—

Vol. 9, No. 4—Design and analysis of industrial experiment: *O. L. Davies* (Dutch text). Proefzetten in de industrie: *H. C. Hamaker.* Tien jaar wiskundige statistiek: *D. van Dantzig.* Tien jaar ontwikkeling van de maatschappelijke statistiek: *J. Idenburg.* Tien jaar ontwikkeling van de industriële statistiek: *J. van Ettinger.*

SWITZERLAND—

Schweizerische Zeitschrift für Volkswirtschaft und Statistik—

March 1956—Die amerikanische Beschäftigungspolitik der dreissiger Jahre im Lichte der Beschäftigungstheorie von Keynes: *B. Fritsch.*

INTERNATIONAL—

International Labour Review—

January 1956—International mobility of manpower in Western Europe: *X. Lannes.*

February 1956—Compensatory public works programmes and full employment: a study of the mobility of resources in the construction industry: *E. J. Howenstein.* International mobility of manpower in Western Europe: II, *X. Lannes.*

World Health Organization—Epidemiological and Vital Statistics Report—

Vol. 9, No. 1—Mortality from accidents (other than transport accidents).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

LIST OF ADDITIONS TO THE LIBRARY.

Since the issue of Part I, 1956, the Society has received the publications enumerated below.

I.—OFFICIAL PUBLICATIONS.

United Kingdom. *Central Office of Information.* The British system of taxation. London, H.M.S.O., 1956. [ii], 45p. 9¼″. 2*s.* 6*d.* (C.O.I. Reference Pamphlet 10)

Local government in Britain. London, H.M.S.O., 1955. vi, 30p. 9¼″. 1*s.* 9*d.* (C.O.I. Reference Pamphlet 1)

Forestry Commission. Report of the Committee on hedgerow and farm timber, 1955. London, H.M.S.O., 1955. 64p. 12 photos. 9¾″. 3*s.* 6*d.*

Ministry of Education. Technical education. London, H.M.S.O., 1956. 43p. 9¾″. 1*s.* 6*d.* (Cmd. 9703)

Ministry of Fuel and Power. Report of the Committee of Inquiry into the electricity supply industry. London, H.M.S.O., 1956. viii, 189p. 2 fold. tables. map. 9¾″. 6*s.* 6*d.* (Cmd. 9672)

Ministry of Health & Secretary of State for Scotland. Report of the Committee of Enquiry into the cost of the National Health Service. London, H.M.S.O., 1956. xii, 309p. maps. 9¾″. 9*s.* (Cmd. 9663)

Ministry of Housing and Local Government. Distribution of rateable values between different classes of property in England and Wales. London, H.M.S.O., 1956. 34p. 9¾″. 1*s.* 6*d.* (Cmd. 9718)

Ministry of Labour and National Service. Cost of Living Advisory Committee. Report on proposals for a new index of retail prices. London, H.M.S.O., 1956. iv, 24p. 9¾″. 1*s.* 6*d.* (Cmd. 9710)

Monopolies and Restrictive Practices Commission. Report on the supply of sand and gravel in central Scotland. London, H.M.S.O., 1956. iv, 109p. 9¾″. 4*s.* 6*d.*

Scotland. General Registry Office. Census 1951. Scotland. Volume IV. Occupations and industries (with some particulars of (*a*) the extent to which people live in one area but work in another, and (*b*) occupations in relation to school leaving age). Edinburgh, H.M.S.O., 1956. xlvi, 658p. 13″. 120*s.*

Select Committee on Nationalised Industries. Special report from the Select Committee on Nationalised Industries: together with the proceedings of the Committee, minutes of evidence and appendices. London, H.M.S.O., 1955. vi, 41p. 9½″. 2*s.* 6*d.* (H.C. 120)

Argentina. *Direccion Nacional del Servicio Estadistico.* IV censo general de la nacion 1947. Tomo I. Censo de poblacion. Buenos Aires, [n.d.] xci, 727p. 10¾″.

Denmark. *Arbejds- og Socialministeriet.* Efterundersøgelse af elever udskrevet af blindeforsorg: en redegørelse ved arbejds- og socialministeriernes statsvidenskabelige konsulent. [An investigation into the living-conditions of blind persons having finished their education in the Danish institutions for the blind.] With an English summary. *Socialt Tidsskrift* (1955). 55p. 9″. *kr.*3. (Økonomisk-statistiske Undersøgelser Nr. 21)

Statistiske Departement. Erhvervsgeografisk materialesamling, folketaellingen 1950. Data of industrial geography from the population census 1950. København, 1956. 207p. map. 9¼″. *kr.* 4 (Statistiske Meddelelser 4.162.6)

Finland. *Statistiska Centralbyrån.* Tulo- ja omaisuustilastoa. Vuodelta 1953 . . . Statistics of income and property in 1953. Helsinki, 1955. Various paging. 11¾″. (Official Statistics of Finland IV: B 20)

Väestötilastoa . . . 1950 population census. Vol. II. Population by industry and industrial status. Helsinki, 1955. 104p. 11¾″. (Official Statistics of Finland VI. C 102).

Germany. *Statistisches Bundesamt.* Organisation und Technik des Volkszählungswerkes 1950. Stuttgart-Köln, 1956. 308p. 11¾″. DM15. (Statistik der Bundesrepublik Deutschland Bd. 31)

Die Untermieten in der Bundesrepublik Deutschland nach der repräsentativen Nacherhebung vom 22.6.1951. Stuttgart-Köln, 1956. 120p. 11¾″. DM6. (S.B.D.Bd. 43)

Greece. *Statistique Nationale de la Grèce.* L'industrie à domicile au recensement du 7 avril 1951. Athènes, 1956. xii, 70p. 13″. *dr.*20

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

International Labour Office. Unemployment insurance schemes. Geneva, 1955. iv, 254p. 9½″. 9*s*. (Studies and Reports N.S. 42)

Israel. *Central Bureau of Statistics* & **Falk Project for Economic Research in Israel.** Provisional estimates of Israel's national expenditure (1952–1953); by Harold Lubell. Jerusalem, 1956. [44]p. Various paging. 12¼″. 600 *pruta*. (Special Series No. 44). English and Hebrew.

Nigeria. *Eastern Region. Information Service.* Eastern Nigeria. Enugu, 1956. London, Office of the Commissioner for Eastern Nigeria, [etc]. 55p. photos. 11″. 3*s*. 6*d*.

Organisation for European Economic Co-operation. The hides and skins industry in Europe: a study on raw hides and skins, the tanning industry and the footwear industry; prepared by the Hides and Skins Committee in December, 1955. Paris, 1955. 141p. 9½″. 10*s*. (Trends in Economic Sectors)

O.E.E.C. Statistical bulletins. Industrial statistics 1900–1955. Paris, 1955. [x], 157p. 10¾″ × 14¼″. 14*s*.

The textile industry in Europe. Paris, 1956. 143*p*. 9½″. 8*s*. (Trends in Economic Sectors)

Productivity measurement. II. Plant level measurements: methods and results; prepared by Mr. Jean Carrié . . . Paris, 1956. 194p. 9½″. 10*s*. (Project No. 235)

Portugal. *Instituto Nacional de Estatistica.* Inquerito as explorações agricolas do continente, 1954. III. Provincias do Minho. Tras-os-Montes e Alto Douro. Lisboa, [1955]. xi, 203p. map. 10¼″. $50.

IX recenseamento geral da população no continente e ilhas adjacentes em 15 de dezembro de 1950. Anexo. Inquerito as condições de habitação da familia. Lisboa, 1954. 233, [6]p. maps. fold. diagr. 10½″. $40.

Spain. *Instituto Nacional de Estadistica.* Estadistica de emisiones de capital año 1954. Madrid, 1955. 15p. 10¾″.

Sweden. *Statistiska Centralbyrån.* Folkräkningen den 31 december 1950. VII. Urvalsundersökningar: statistiken över inkomst. Census of the population in 1950. VII. Sample surveys: statistics of income. Stockholm, 1956. xvi, 78°, 93p. 9½″. (Sveriges Officiella Statistik)

Switzerland. *Eidgenössisches Statistisches Amt.* Schweizerische Volkssterbetafeln 1941/50 und 1948/53: Grundzahlen und Nettowerte. Bern, 1955. 23*, 85p. 11¾″. (Statistische Quellenwerke der Schweiz Heft 282)

Union of South Africa. *Bureau of Census and Statistics.* Population census, 7th May, 1946. Vol. V. Occupations and industries of the European, Asiatic, coloured and native population. Pretoria, 1955. x, 219p. 12¼″. 47*s*. 6*d*.

United Nations. *Department of Economic and Social Affairs.* Proceedings of the World Population Conference, Rome, 31 August–10 September 1954. Summary report. New York, 1955. London, H.M.S.O. iv, 207p. 11″. 7*s*. 6*d*.

Department of Economic and Social Affairs. Population Branch. Age and sex patterns of mortality: model life-tables for under-developed countries. New York, 1955. [iv], 38p. 11″. 3*s*. (Popul. Stud. No. 22)

Manuals on methods of estimating population. Manual II. Methods of appraisal of quality of basic data for population estimates. New York, 1955. viii, 67p. map. 11″. 5*s*. (Popul. Stud. No. 23)

Department of Public Information. Yearbook of the United Nations, 1954. New York, 1955. x,[ii], 656p. fold. col. diagrs. 10″. 70*s*.

Economic Commission for Europe. Economic survey of Europe in 1955: including studies of investment problems and policies of European countries and labour market problems in Western Europe; prepared by the Research and Planning Division, Economic Commission for Europe. Geneva, 1956. [352]p. Various paging. 11″. 18*s*.

The European housing situation: prepared by the secretariat of the Economic Commission for Europe. Geneva, 1956. 56p. 11″. 3*s*. 6*d*.

Economic Commission for Europe & Food and Agriculture Organization. Output and expenses of agriculture in some European countries. Second report: 1950–1953. Geneva, 1955. [iii], 230p. 11¼″. 5*s*.

Educational, Scientific and Cultural Organization. Film and cinema statistics: a preliminary report on methodology with tables giving current statistics. Paris, [1955]. 111p. 11″. 3*s*. (U.N.E.S.C.O. Statistical Reports and Studies)

Food and Agriculture Organization. Report of the third special meeting on the economic aspects of the rice industry, Bangkok, Thailand, 30 September–7 October 1955. Rome, 1955. v, 49p. 9″. 2*s*. 6*d*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Vienna. *Statistisches Amt.* Historisch-statistische Übersichten von Wien; zusammengestellt von Dr. Felix Olegnik. I. Teil: Naturverhältnisse, Gebiet, Bevölkerung, Gesundheits- und Wohlfahrtswesen. Wien, Magistrat der Stadt, 1956. 150p. 11½″. (*Mitteilungen aus Statistik und Verwaltung der Stadt Wien*, 1956, Sonderheft 1)

West Bengal. *State Statistical Bureau.* Economic survey of small industries, 1954. District—Maldah. [iv], 243p. District—West Dinajpur. [iv], 255p. 15s. Alipore, Superintendent, Government Printing, 1955. 9¾″.

West Sussex. *County Council. County Education Committee.* Road safety: road accidents to children under sixteen in West Sussex, 1954. Chichester, 1955. 10p. 4 diagrs. 13″.

Yugoslavia. *Federal Statistical Office.* Konačni rezultati popisa stanovništva od 15 marta 1948 godine, Knjiga X. Domačinstva. Final results of the population census of March 15th, 1948. Vol. X. Households. Beograd, 1956. lix, 188p. 11½″.

II.—AUTHORS AND MISCELLANEOUS.

Angelopoulos, Angelos. L'atome unira-t-il le monde? aspects économiques, sociaux, politiques. Paris, Pichon & R. Durand-Auzias, 1956. [iv], 216p. 8¼″. *fr.* 750.

Bennett, M. K. The food economy of the New England Indians, 1605–75. *J. Polit. Econ.* (1955), **63**, 369–397. 9½″.

Berkson, Joseph. Estimate of the integrated normal curve by minimum normit chi-square with particular reference to bio-assay. *J. Amer. Statist. Ass.* (1955), **50**, 529–549. 9¼″.

Bethlem Royal Hospital & Maudsley Hospital. Triennial statistical report: years 1949–1951; by C. P. Blacker, A. T. Gore. London, 1955. xii, 115p. 8½″. 10*s.* 6*d.* The first of a series of triennial reports.

Blackwell, David. On transient Markov processes with a countable number of states and stationary transition probabilities. *Ann. Math. Statist.* (1955), **26**, 654–658. 10″.

Bok, Derek Curtis. The first three years of the Schuman plan. Princeton, University, International Finance Section, 1955. [vi], 79p. 9″. 25*c.* (Stud. Int. Finance. 5)

Bologna, University. L'Istituto e la Scuola di Statistica dell' Università di Bologna. [Bologna, 1955?] 15p. 15 plates, photos. 8¾″.

Bristol and Bath Productivity Association. "What about work study?" An enquiry into its benefits and snags; by a Bristol Working Party. Bristol, The Association, 1956. 40p. illus. 8½″. 3*s.* Also obtainable from the British Productivity Council.

Brunsviga Maschinenwerke AG. Directions for use of the Brunsviga 20 calculating machine ... Braunschweig, Brunsviga Maschinenwerke AG. [1952]. 32p. fold. diagr. 5¾ × 8″.

Cambridge. University. *Department of Applied Economics.* Input-output and the social accounts; Richard Stone. Reprinted from *The Structural interdependence of the economy*, University of Pisa, 1955. 155–172. 9½″. (Repr. Ser. 102)

Serial correlation in regression analysis I; by G. S. Watson. *Biometrika* (1955), **42**, 327–341. 10¾″. (Repr. Ser. 103)

Freeing Europe's trade and payments; by W. B. Reddaway. *Times Rev. Ind., Lond. & Camb. Econ. Bull.* (1955), No. 15. 7p. 9¾″. (Repr. Ser. 104)

Central Mortgage and Housing Corporation. Housing and urban growth in Canada; a brief from Central Mortgage and Housing Corporation to the Royal Commission on Canada's Economic Prospects. Ottawa, 1956. 35p. 9¾″.

Centre National de la Recherche Scientifique. Les modèles économétriques. Paris, C.N.R.S., 1955. 147p. 9½″. *fr.* 750. (Cahiers du Séminaire d'Econométrie No. 3). Contributions by P. Maillet, J. Tinbergen, E. Malinvaud, H. Wold, J. Marschak and R. Janin.

Ceppellini, R., Siniscalco, M. & Smith, C. A. B. The estimation of gene frequencies in a random-mating population. *Ann. Hum. Genet.* (1955), **20**, 97–115. 10½″.

Chartier, F. Les sondages à plusieurs degrés. [Paris], Institut National de la Statistique et des Etudes Economiques, Ecole d'Application, 1955. [iv], 8, 96p. 10½″. Typescript.

Cheeseman, E. A. & Stevenson, A. C. Records of deaf mutism in Northern Ireland. *Brit. J. Prev. Soc. Med.* (1955), **9**, 196–200. 9¾″.

Cohen, John & Hansel, Mark. Risk and gambling: the study of subjective probability. London, Longmans, Green, 1956. x, 153p. 7½″. 14*s.*

Colombo, Bernardo. Sul rapporto dei sessi nelle nascite e nei concepimenti. Padova, Cedam, 1955. [ii], 93p. 10″. *L.*1000.

Co-operative Union, Ltd. Co-operative shops, 1955: a census of retail outlets and main services operated by co-operative societies. Manchester, Co-operative Union, 23p. 8¾″ × 10½″. 1*s.*

Coward, K. H. & Irwin, J. O. The second international standard for vitamin D_3. *Bull. Wld. Hlth. Org.* (1954), **10**, 875–894. 9½″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Davidson, Andrew Rutherford. The history of the Faculty of Actuaries in Scotland, 1856–1956. Edinburgh, The Faculty of Actuaries, 1956. xviii, 291p. 26 plates (2 fold.). 8½″. 20*s*.

Davies, R. Llewelyn, Nightingale, J. W. & Bailey, N. T. J. Survey of space and services requirements in two agricultural research laboratories. *Nature* (1955), **176**, 999–1001. 5p. 8½″.

Duon, Gaston. De la théorie à la pratique des indices statistiques: application aux problèmes économiques, sociaux et démographiques. Paris, Eyrolles & Gauthier-Villars, 1955. 156p. 9½″. *fr*. 1300. English summary.

Economist. *Intelligence Unit*. The businessman's guide to Britain, 1956. London, Chatto & Windus, [1956]. ix, 155p. maps. 8″. 10*s*. 6*d*.

Federer, Walter T. Experimental design: theory and application. New York, Macmillan, [1956]. xix, [3], 544, 47p. 9¼″. 77*s*.

Fieller, E. C. & Smith, C. A. B. Note on the analysis of variance and intraclass correlation. *Ann. Eugen.* (1951), **16**, 97–104. 10¾″.

Finney, D. J. & Outhwaite, Anne D. Serially balanced sequences. *Nature* (1955), **176**, 748. 2p. 8½″.

Finney, D. J. & Varley, G. C. An example of the truncated Poisson distribution. *Biometrics* (1955), **11**, 387–394. 9¼″.

Fitzpatrick, Paul J. The early teaching of statistics in American colleges and universities. *Amer. Statistician* (1955), **9**, 12–18. 11″.

Gadgil, D. R. & Dandekar, V. M. Primary education in Satara District; reports of two investigations. Poona, Gokhale Institute of Politics and Economics, 1955. vi, 172p. 9½″. 9*s*. (Gokhale Inst. Polit. Econ. Publication 32)

Gehrels, Franz & Johnston, Bruce F. The economic gains of European integration. *J. Polit. Econ.* (1955), **63**, 275–292. 9½″.

Glover, J., Robinson, P. & Taylor, J. Assessing the reliability of rainfall if monthly falls are not independent. *J. Agric. Sci.* (1955), **46**, 387–388. 10½″.

Good, I. J. Conjectures concerning the Mersenne numbers. *Math. Tab., Wash.* (1955), **9**, 120–121. 10¾″.

The likelihood ratio test for Markoff chains. *Biometrika* (1955), **42**, 531–533. 10¾″.

Some terminology and notation in information theory. London, Institution of Electrical Engineers, 1955. 5p. 11″. (Monograph No. 155R). To be republished in Part C of the *Proceedings* of the Institution.

Grünhut, Max. Juvenile offenders before the courts. Oxford, Clarendon Press, 1956. vii, 143p. 8½″. 21*s*.

Hamaker, H. C. Experimental design in industry. *Biometrics* (1955), **11**, 257–286. Contribution to the 3rd International Biometric Conference held in Bellagio, 1–5 September 1953.

Hamaker, H. C. & Strik, R. van. The efficiency of double sampling for attributes. *J. Amer. Statist. Ass.* (1955), **50**, 830–849. 9¼″.

Hamilton, W. R. The principles of modern book-keeping; by W. R. Hamilton. 3rd ed. by B. J. Davis. London, Gee & Co., 1956. 109p. 8½″. 10*s*.

Hebden, Jack & Jowett, Geoffrey H. The accuracy of sampling coal. *Appl. Statist.* (1952), **1**, 179–191. 2 plates. 10″.

Heilbroner, Robert L. The great economists: their lives and their conceptions of the world: by Robert L. Heilbroner. Revised for publication in England with two supplementary chapters by Paul Streeten. London, Eyre & Spottiswoode, 1955. 320p. 8½″. 28*s*.

Hodges, J. L., *Jr*. On the noncentral beta-distribution. *Ann. Math. Statist.* (1955), **26**, 648–653. 10″.

Institute of Chartered Accountants of Scotland. Examination and training of apprentices: report of the Special Committee. Edinburgh, The Institute, 1956. 36p. 13¼″.

Presidential address; by Sir Ian F. Bolton ... Delivered ... on 28th March, 1956. Edinburgh, The Institute, 1956. 4p. 8¾″.

International Association for Research in Income and Wealth. Bibliography on income and wealth. Vol. IV. 1951; edited by Phyllis Deane. Cambridge, Bowes & Bowes, 1954. 82p. 9¾″. 37*s*. 6*d*.

International Statistical Institute. International Statistical Conferences. December 1951. India. *Bull. Int. Statist. Inst.*, **33**, Part 1. ii, 143p. 9″. $1.25.

Irwin, J. O. A distribution arising in the study of infectious diseases. *Biometrika* (1954), **41**, 266–268. 10¾″.

Statistik in der medizinischen Forschung. *Industr. Organisation* (1954), **23**, 395–402. 11¾″.

Irwin, J. O. & Standfast, A. F. B. Litter-mate assays of pertussis vaccine. *J. Hyg. Camb.* (1955), **43**, 106–111. 10″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

James, G. S. Cumulants of a transformed variate. *Biometrika* (1955), **42**, 529–531. 10¾″.

Jones, William O. [Review of] *Colonial Social Accounting* [by Phyllis Deane]. *J. Amer. Statist. Ass.* (1955), **50**, 665–676. 9¼″.

Jowett, Geoffrey H. The accuracy of systematic sampling from conveyor belts. *Appl. Statist.* (1952), **1**, 50–59. 10″.

The comparison of means for industrial time series. *R. Statist. Soc. Conf. Industr. Appl. Statist.* (1953). 6p. [3]l. diagrs. 13″.

Multiple regression between variables measured at the points of a plane sheet. *Appl. Statist.* (1955), **4**, 80–89. 10″.

Sampling properties of local statistics in stationary stochastic series. *Biometrika* (1955), **42**, 160–169. 10¾″.

Kilpatrick, S. J., Mathers, J. D. & Stevenson, A. C. The importance of population fertility and consanguinity data being available in medico-social studies: some data on consanguineous marriages in Northern Ireland. *Ulster Med. J.* (1955), **24**, 113–122. 9½″.

Kuczynski, Jürgen. Die Geschichte der Lage der Arbeiter unter dem Kapitalismus. General-register; [von] Ruth Hoppe. Berlin, Tribüne, 1956. 265p. 8¼″.

Lenti, Libero. Problemi economici d'oggi. Milano, Garzanti, 1956. 134p. 7¼″. *L*300.

London Transport Executive. Health in industry: a contribution to the study of sickness absence, experience in London Transport. London, Butterworth, on behalf of London Transport Executive, 1956. [v], 177p. 8¾″. 35*s*.

Malzberg, Benjamin & Lee, Everett S. Migration and mental disease: a study of first admissions to hospitals for mental disease, New York, 1939–1941. New York, Social Science Research Council, 1956. x, 142p. 9″. $1.50.

Meade, James E. The Belgium-Luxembourg Economic Union, 1921–1939: lessons from an early experiment. Princeton, University, International Finance Section, 1956. [ii], 41p. 9″. gratis. (Essays Int. Finance, 25)

Mehta, M. M. Structure of Indian industries (a statistical study in the size, location and integration of industrial units in the seven selected industries of India, viz., cotton, jute, sugar, iron and steel, coal, paper and cement) ... Bombay, Popular Book Depot, 1955. xix, 340p. 9¾″. *Rs.* 22/8.

Midzuno, H., Hayashi, C., Matusita, K. & Aoyama, H. [Tables for nonparametric tests. Tokyo, Asakura Publishing Co., 1954] [x], 229p. 8¼″. In Japanese.

Mills, Frederick C. Statistical methods ... 3rd ed. London, Pitman, 1956. xviii, 842p. 9¼″. 50*s*. Third edition of *Statistical methods applied to economics and business*.

Morton, J. E. Urban mortgage lending: comparative markets and experience. Princeton, University Press, 1956. London, Oxford University Press. xxi, 187p. 9″. $4. (N.B.E.R. Studies in Urban Mortgage Financing, 6)

Murray, Keith A. H. Agriculture. London, H.M.S.O. & Longmans, Green, 1955. xii, 422p. 9½″. 30*s*. (History of the Second World War, U.K. Civil Series)

Nash, Stanley W. Contribution to the theory of experiments with many treatments. Berkeley & Los Angeles, University of California Press, 1956. London, Cambridge University Press. [ii], 167–183p. 10¼″. 25*c*. (University of California Publications in Statistics, Vol. 2, No. 8)

National Bureau of Economic Research. Capital formation and economic growth: a conference of the Universities-National Bureau Committee for Economic Research. Princeton, University Press, 1955. xiii, 677p. 9″. $12. (Special Conference Series No. 6)

Distribution's place in the American economy since 1869; by Harold Barger. Princeton, University Press, 1955. London, Oxford University Press. xviii, 222p. 9″. $4.50. (N.B.E.R. No. 58)

Interest as a source of personal income and tax revenue; Lawrence H. Seltzer. New York, N.B.E.R., 1955. [v], 1248–1330p. 9¼″. (N.B.E.R. Occ. Pap. 51). Reprinted from *J. Amer. Statist. Ass.* (1955), **50**.

Personal income during business cycles; by Daniel Creamer, with the assistance of Martin Bernstein. Princeton, University Press, 1956. London, Cumberlege. xlii, 166p. 9″. $4. (Studies in Business Cycles 6)

Policies to combat depression: a conference of the Universities-National Bureau Committee for Economic Research. Princeton, University Press, 1956. London, Cumberlege. x, 417p. 9″. $8.50. (Special Conference Series 7)

National Institute of Economic and Social Research. A classified list of large companies engaged in British industry. London, 1955. 39p. 11″. 5*s*. For private circulation.

The measure of income for shareholders and for taxation; by S. J. Prais. *Accounting Res.* (1955), **6**, 187–201. 15p. 9¾″. (Repr. Ser. 5)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

National income and expenditure of the United Kingdom, 1870–1952; by James B. Jefferys and Dorothy Walters. *Income and Wealth Ser. V.* (1956). 40p. 8½″. (Repr. Ser. 6)

Netherlands. *Centraal Planbureau.* Activiteitsanalyse, lineaire programmering en hun toepassing; door S. L. Louwes. *De Economist* (1955), **103**, 817–853, (1956), **104**, 40–56. 53p. 9½″. (Repr. Ser. 44)

Loonverschillen tussen industriële bedrijfstakken 1920–39 en 1947–53; door J. S. Cramer. *Econ.-statist. Ber.* (1956), **41**, 31–33. 8p. 9½″. (Repr. Ser. 45)

Neville, E. H. Rectangular-polar conversion tables; designed and compiled by E. H. Neville. Cambridge, published for the Royal Society at the University Press, 1956. xxxiii, 109p. 11″. 30*s.* (Royal Society Math. Tables, 2)

Nichols, P. J. R. & Bailey, N. T. J. The accuracy of measuring leg-length differences: an "observer error" experiment. *Brit. Med. J.* (1955), **2**, 1247. 4p. 8½″.

Nielsen, Kaj L. Methods in numerical analysis. New York, Macmillan, 1956. xiii, 382p. 8¼″. 48*s.* 6*d.*

North Carolina. University. *Institute of Statistics.* Cenotypic and phenotypic correlations in soybeans and their implications in selection; by H. W. Johnson, H. F. Robinson and R. E. Comstock. *Agron. J.* (1955), **47**, 477–483. 11″. (Repr. Ser. 74)

Simultaneous tests of linear hypotheses; by M. N. Ghosh ... *Biometrika* (1955), **42**, 441–449. 10½″. (Repr. Ser. 75)

Estimation of the mean and standard deviation by order statistics. Part III; A. E. Sarhan ... *Ann. Math. Statist.* (1955), **26**, 576–592. 9¾″. (Repr. Ser. 77)

Effect of raw-material ratios on absorption of whiteware compositions; by W. C. Hackler, W. W. Kriegel and R. J. Hader ... *J. Amer. Ceram. Soc.* (1956), **39**, 20–25. 11″. (Repr. Ser. 78)

Outhwaite, Anne D. & Rutherford, A. Covariance analysis as an alternative to stratification in the control of gradients. *Biometrics* (1955), **11**, 431–440. 9¼″.

Pedestrians' Association for Road Safety. The investigation of accidents: statistics that mislead. *Pedestrian* (1956), Spring, 7–12. 7¼″.

Rexford-Welch, S. C., *editor.* The Royal Air Force medical services. Vol. II. Commands. London, H.M.S.O., 1955. xxiv, 703p. illus. 9½″. 75*s.* (History of the Second World War, U.K. Medical Series)

Rich, Earl R. Egg cannibalism and fecundity in *tribolium. Ecology* (1956), **37**, 109–120. 10¾″.

Roy, Purnendu Mohon. On the relation between B.I.B. and P.B.I.B. designs. *J. Indian Soc. Agric. Statist.* (1954), **6**, 30–47. 9¾″.

Scott, J. D. & Hughes, Richard. The administration of war production. London, H.M.S.O. & Longmans, Green, 1955. xii, 544p. 9½″. 37*s.* 6*d.* (History of the Second World War, U.K. Civil Series)

Smith, C. A. B. The linear function maximizing intraclass correlation. *Ann. Eugen.* (1953), **17**, 286–292. 10¾″.

The separation of the sexes of parents in the detection of linkage in man. *Ann. Eugen.* (1954), **18**, 278–301. 10½″.

A simplified heterogeneity test. *Ann. Eugen.* (1952), **17**, 35–36. 10¾″.

Snow, John. Report of a meeting held at the London School of Hygiene and Tropical Medicine to celebrate the centenary of the publication of John Snow's classic treatise "On the mode of communication of cholera". Snow—the man and his times; by J. M. Mackintosh. Snow—an appreciation; by A. Bradford Hill. *Proc. Roy. Soc. Med.* (1955), **48**, 1004–1012. 9p. illus. 9¾″.

Somogyi, Stefano. L'heure de la naissance et du décès. *Bull. Int. Statist. Inst.* (1954), **34**, liv. 3. 27p. 9½″.

Primi resultati della rilevazione generale sulla morbosità ospedaliera in Italia con particolare riferimento alla Sicilia. *Riv. ital. Econ. Demogr. Statist.* (1954), **8**, n.ri 3–4. 27p. 9¾″.

Prospettive demografiche per l'Italia 1950–1970. *Difesa soc.* (1955), **34**, vol. 3. 46p. 9½″.

Sistematica della misurazione statistica della morbosità. *Previdenza soc.* (1955), sett.-ott. 32p. 9½″.

Sprowls, R. Clay. Elementary statistics for students of social science and business. New York, London, [etc.], McGraw-Hill, 1955. xiii, 392p. 9″. 41*s.* 6*d.*

Stevenson, A. C. & Cheeseman, E. A. Hereditary deaf mutism, with particular reference to Northern Ireland. *Ann. Hum. Genet.* (1956), **20**, 177–231. 11″.

Titmuss, R. M. The social division of welfare. Liverpool, University Press, 1956. 23p. 8½″. 2*s.* 6*d.* (6th Eleanor Rathbone Memorial Lecture, Dec. 1st, 1955)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Uppsala. University. *Institute of Statistics.* Selected publications. Vol. 11. [Uppsala, 1956.] Various paging. 9″.

Vajda, Stefan. The theory of games and linear programming. London, Methuen, 1956. New York, John Wiley. vi, 106p. 6½″. 8*s.* 6*d.* (Methuen's Monographs on Physical Subjects)

Waasdijk, T. van. Hire purchase credit in South Africa: an economic survey. Johannesburg, Witwatersrand University Press, 1956. xvi, 174p. 8½″. 21*s.*

Wallis, W. Allen & Roberts, Harry V. Statistics: a new approach. Glencoe, Illinois, The Free Press, 1956. xxxviii, 646p. 9¼″. $6.

West, K. Hall. Accounting practice. London, Gee & Co., 1956. 132p. 8½″. 21*s.*

West Indies. University College of the West Indies. *Institute of Social and Economic Research.* Research programme and progress report. Jamaica, 1955. [ii], 29p. 6 photos. 9½″.

World oilmen: the story in pictures of the men who provide the world's oil; text by A. M. Melland. London, Sylvan Press, 1955. [100]p. photos. 12″.

Zwiedineck-Südenhorst, Otto von. Mensch und Wirtschaft: Aufsätze und Abhandlungen zur Wirtschaftstheorie und Wirtschaftspolitik. 1. Band. Berlin, Duncker & Humblot, 1955. ix, [3], 440p. 9¼″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Vol. 119.] [Part IV

Journal of the Royal Statistical Society

SERIES A (GENERAL)

PART IV, 1956.

Estimating Expenditure, Saving and Income from Household Budgets

By Dorothy Cole and J. E. G. Utting

[Read before the Royal Statistical Society, May 16th, 1956, the President, Professor E. S. Pearson, C.B.E., in the Chair]

Introduction

The phenomenon of expenditure exceeding income in family budget studies has a long history. At the end of the eighteenth century Sir Frederick Morton Eden (ed. Rodgers, 1928), one of the earliest budget investigators in this country, found quite frequently that the expenditure of the families he interviewed exceeded their stated income. More recently and more significantly the pre-war middle class budget inquiry organized by the Civil Service Statistical and Research Bureau (Massey, 1942, and Prais and Houthakker, 1955) showed an excess of recorded expenditure over recorded income for all income groups, whilst the 1950 survey of consumers' expenditure carried out by the U.S. Bureau of Labor Statistics found that, although for some individual families income exceeded expenditure plus saving, in general a negative balance predominated. Other cases could be quoted. It need hardly be said that such a lack of balance has important implications for users of budget data. If, for instance, the explanation of the lack of balance is that expenditure is overstated, it would be important to know not only that but also whether it is overstated equally as between different income groups or as between commodity groups. If income is understated it would be extremely important to know whether such understatement applies to all income groups equally or to some more than others.

The difficulty in appraising budgetary data along these lines has hitherto been that very few studies have in fact collected enough information to arrive at a complete balance. Saving is the element completing the identity $I = C + S$. Excess of expenditure over income could mean that households were dis-saving. This is usually unlikely on a large scale, but without savings information the picture is not complete. The U.S. Bureau of Labor Statistics inquiry referred to above is particularly interesting because it did in fact collect income, saving and expenditure information. The Social Accounts of Cambridgeshire inquiry* carried out by the Department of Applied Economics is one of the few surveys in this country which has tried to collect all three items, defining them in such a way as to cover all monetary receipts and disbursements by private individuals. Our experience has been the same as that of other surveys. Estimated mean expenditure and saving of the sample households exceeds income by £104.

We shall attempt in this paper to collect together the evidence provided by our data on the questions posed above, but we first present a brief description of the survey and some of its principal results.

The Survey Methods†

Our survey was conducted in the administrative county of Cambridgeshire, using a stratified random sample of about 4,000 addresses drawn from local rating lists. The strata and sampling fractions were:

* Henceforth this survey will be referred to as the S.A.C. inquiry.

† For a more detailed description see Utting and Cole (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Stratum	*Description*	*Sampling fraction*
1	Urban area, rateable value £1 to £30	6·7 per cent.
2	Urban area, rateable value £31 or more	13·3 ,,
3	Rural area, rateable value £1 to £13	6·7 ,,
4	Rural area, rateable value £14 or more	20·0 ,,

The residents at the selected addresses were grouped into households and an attempt was made to obtain information relating to every *household* at each address, by interviewing every *individual* aged 15 or more and not still at school. All information was collected by interview; no special records were kept, but the respondents sometimes referred to bills, cheque stubs, etc. Interviewing was spread as evenly as possible over the year April, 1953 to April, 1954.

Information about expenditure was sought from the whole sample. This related to all expenditure during the preceding seven days and also, for particular items, during the preceding four weeks, quarter or year. In addition, one half of the sample was asked for information about income and the other half for information about changes in asset holdings and also for a broad income group. There is thus no balance of expenditure, saving and income for individual households, but only for our estimates for broad groups of households. Table 1 summarizes the various time periods covered for different groups of items.

TABLE 1

Periods Covered in the S.A.C. Inquiry

Item	*Period*	*Note*
Income . . .	Year	1. Wages, salaries, government transfers, rent, dividends and interest: Interviewers recorded amount received "last time" from which a year's figure was calculated. In some cases the figure would be a single amount received in the last year. 2. Self-employment income: The amount of business profit for the latest available accounting period.
Saving . . .	Year	1. Contractual saving (life insurance and mortgage repayments): Interviewers recorded the amount paid "last time" from which a year's figure was calculated. 2. Other saving: Interviewers recorded the change in the holding of assets as between the date of interview and a year earlier.
Expenditure . .	Week	1. All expenditure in the seven days preceding the day of interview was recorded.
		In addition:
	Month	1. Clothing, household textiles, etc., were covered for four weeks preceding the day of interview.
	Year	1. A list of major items (major durables) was covered for twelve months preceding the day of interview. 2. Rent, rates, etc., were recorded as the amount paid last time and a year's figure calculated.
	Quarter	1. Gas and electricity (other than slot meters): Interviewers recorded the amount of the last bill paid. The same technique was used for telephone bills, only these are six-monthly. 2. Coal: Interviewers recorded the value of the amount delivered in the thirteen weeks preceding the day of interview.

For purposes of analysis we took the *household* as our basic unit and used all households for which we had complete, or nearly complete, expenditure information. Using this criterion, the overall response rate was 75 per cent., giving a sample of 3,152 households divided up as follows:

Stratum	*Income forms*	*Saving forms*	*Total*
1	640	665	1,305
2	164	169	333
3	603	587	1,190
4	148	176	324

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

For estimation purposes we combined all the figures for "last week's" expenditure for each household into broad commodity groups and so obtained estimates of mean weekly expenditure for each group and hence for all items combined. The longer period information could not be combined in quite this way since the length of period for different items varied even within a commodity group. We therefore made separate estimates for groups of items which were covered for the same period and converted these to annual rates before combining them into commodity groups. There are thus two different sets of expenditure estimates which will be referred to as "last week's" and "period" expenditure respectively. Although last week's total expenditure can be estimated directly from household totals, a period total can only be built up from the parts. Similar methods were used for the estimation of income and saving; total regular income can be estimated independently of the parts and so can total saving, all the figures for which related to the past year.

For most items of expenditure, saving and income, there are some households in our sample for which the amount is not known. In estimating expenditure means we have just ignored these cases, equivalent to assuming that their mean expenditure is the same as that of the known cases including zeros. For income and saving, however, we have assumed that the unknown cases are definitely non-zero and that their mean is the same as that of the known non-zero cases. We have been influenced in this decision mainly by the nature of our questionnaire and the responses it elicited, but also by the fact that the latter method gives a closer correspondence between the sum of the parts of income or saving and the independent estimate of the sum. The estimated means and their sampling errors are obtained from the normal formulae, with appropriate adjustment for income and saving. When a period estimate has been built up from several parts, we have added the variances as though those parts were independent, although this is clearly not the

TABLE 2

Estimated Mean Annual Expenditure of Cambridgeshire Households
1953–54

Commodity Group (for further details see Appendix)	*"Last Week's" Estimate*		*"Period" Estimate*	
	Amount £	*Std. Error* %	*Amount* £	*Std. Error* %
Food	183·0	1	—	—
Alcoholic drink	20·7	4	—	—
Tobacco	31·2	2	—	—
Housing (rent, rates, mortgage interest, repairs)*	49·5	11	52·3	6
Gas	8·4	8	7·9	4
Electricity	11·0	8	8·7	2
Coal and coke	20·8	5	13·3	2
Other fuel	3·3	8	—	—
Telephone, postage, etc.	5·8	8	5·3	3
Licences, etc.	2·6	9	2·8	5
Books and newspapers	9·9	3	10·0	3
Entertainment	15·0	5	—	—
Clothing	55·1	5	71·9	3
Travel	38·6	5	39·8	4
Major durable goods (net)†	83·8	25	35·9	4
Other durable goods	23·3	8	29·2	5
Soaps, polishes, etc.	12·9	3	—	—
Domestic service	8·5	6	9·6	8
Property insurance	2·7	20	2·0	3
Subscriptions, gifts, etc.	13·5	10	13·2	5
N.H.S. charges	1·0	18	1·0	10
Other goods	30·8	4	32·5	3
Other services	40·3	9	46·9	4
Total expenditure	674·3	4	—	—
(Total of items)	(671·7)	—	(648·4)	—

* Excludes £3·5 regular repayment of mortgage principal, which is included in *saving*.
† After deducting £5·9 sales of second-hand items.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

case. The arrangement of our material does not readily permit the calculation of covariances, though it appears that in most of our groups some will be positive and some negative. Our "period" sampling errors are therefore only a rough guide to the order of magnitude.

To make some allowance for the fact that our effective response rate varied somewhat during the year, which may be important if there are genuine seasonal variations in the figures, we divided our sample into four quarters according to the date of interview, and our overall estimates are the unweighted means of independent quarterly estimates. The estimates which are subdivided by household income groups, in which the sample numbers are smaller, have been calculated by the more straightforward method, from all households within the group irrespective of date of interview.

The Principal Results

In Tables 2, 3 and 4 we present overall estimates of household* expenditure, income and saving, and their estimated sampling errors:

It will be seen that for many items of expenditure there is a marked difference between the "last week" estimates and the "period" estimates. The result of varying the length of period is,

TABLE 3

Estimated Mean Annual Saving of Cambridgeshire Households
1953–54

Form of Saving	*Amount* £	*Std. Error*† %
Additions to liquid assets (P.O. savings bank, savings certificates, defence bonds, etc.)	−7·1	80
Net purchases of stocks, shares, annuities	13·2	80
Business saving of self-employed	8·5	33
Increase in bank balance or cash	−12·5	40
Life assurance premiums	21·0	3
Net repayment of housing debt (mortgages, etc.)	−8·8	50
,, ,, of hire purchase debt	−2·2	40
,, ,, of other debt	0·4	1300
,, purchase of houses	25·0	20
less Lump sum receipts (gifts, legacies, matured endowment assurance, etc.)	−30·7	45
Total saving	9·5	60
(Total of items)	(6·8)	

† The estimated sampling errors are very approximate.

TABLE 4

Estimated Mean Annual Income of Cambridgeshire Households
1953–54

Type of Income	*Amount* £	*Std. Error* %
Regular income from main employment	433·2	2
,, ,, ,, other employment	7·7	14
,, ,, ,, self-employment	64·4	17
Rent, dividends and interest	39·7	13
Regular income from govt. pensions, family allowances, etc.	39·9	4
Other regular income	20·6	9
Total regular income	603·3	2
(Sum of regular items)	(605·5)	
Casual income from employment	7·3	16
,, ,, from self-employment	0·1	38
,, receipts of national insurance benefits, etc.	1·7	18
Other casual income	0·7	41
Income tax paid	45·7	8
,, ,, refunded	0·3	40
National Insurance contributions	14·8	2

* The average household size is 2·84 persons.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in fact, even more marked than Table 2 suggests, because many of the commodity groups shown there include particular items which were only covered for the last week. These have been combined with other items, covered for a longer period, to give the period estimate. Table 5 gives a comparison of the week and period estimates after excluding those parts which are common to both.

There is therefore a very real problem of choice between them. With one exception—domestic service—the estimated sampling errors of the period estimates are smaller than those of the last week's estimates. This was, of course, one of the main objects of using the longer periods, and if the estimates were otherwise consistent we might well use an appropriate combination of the two. But in a number of cases (e.g. coal and coke, clothing) the estimated sampling errors are such that it seems very unlikely that the differences are due only to sampling fluctuations. Many factors of a non-sampling character affect the reliability of such estimates, and it seems reasonable that they should affect the two different sets of estimates in different ways. Among such factors are the straightforward one of memory—that some items of expenditure may be completely forgotten—and the more complex one of end-period effect, which leads to items which really belong to an earlier period being carried into the period of account. The combined result of these two factors is difficult to forecast and has not, as yet, been very fully investigated in the field, although Mahalanobis and Sen (1954, p. 9) have described an experiment in India. We are led therefore to rely on the very general conclusions which guided us in our original choice of periods and to say that, in the absence of any positive evidence to the contrary, we prefer the period estimate and have used that for all items for which it exists.*

TABLE 5

Comparison of Week's and Period Estimates
(free of duplication)

Item	*Estimated Annual Expenditure*		*Period Covered*
	Week's Estimate £	*Period Estimate* £	
Paint, distemper, etc.	3·7	5·5	Four weeks
Clothing, other than haberdashery	53·4	70·2	,,
National Health Service charges	1·0	1·0	,,
Minor household durable goods	23·3	29·2	,,
Gas (quarterly account)	5·0	4·5	Quarter
Electricity (quarterly account)	10·2	7·8	,,
Coal and coke	20·8	13·3	,,
Telephone bill	2·0	1·4	Half year
Rates	8·2	7·2	,,
House purchase	8·6	38·7	Year
Licences and stamp duty	1·7	1·8	,,
Major durable goods	89·7	41·8	,,
Magazine subscriptions	0·4	0·4	Annual rate
Subscriptions to clubs, societies, etc.	5·7	5·4	,,
Domestic service	6·2	7·3	,,
Life insurance	27·7	21·0	,,
Property insurance	2·7	2·0	,,

If we use the period estimates of expenditure, we have of necessity to use an estimate of total expenditure which is the sum of the separate items, and we have no direct estimate of its sampling error. It seems reasonable, however, to assume that in percentage terms it does not exceed the estimated error for the week's total expenditure, and we have used that figure wherever a sampling error for total expenditure is quoted. We still have a choice between the direct estimates of total regular income or total saving and the sum of the items of which they are composed.† Here a statistical adjustment procedure would be perfectly possible but again, because of the importance

* This choice is also supported by the comparison with other sources, described below (p. 380).

† We have figures for *total* saving and *total* regular income for 92 per cent. of the households in the appropriate halves of the sample.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of non-sampling errors, we have preferred to make a simple choice, and have used the direct estimate of the sum (the difference is negligible: £603·3 compared with £605·5 for regular income and £9·5 compared with £6·8 for saving). We have chosen this estimate because in the case of saving, where positive and negative items frequently offset one another, it is derived directly from the net balance for each individual household for which this information is complete, rather than as a balance of estimates which include incomplete information. The sampling error of these estimates of regular income and saving can also be estimated directly; we have assumed the same proportionate error for the total of income, after adding the small amounts of irregular income.

We can now strike a balance between our estimates of expenditure, saving and income, as follows:

	£	*Estimated Std. Error* %
1. Expenditure	647·4	4
2. Income tax and national insurance	60·2	7
3. Saving	9·5	60
4. 1 + 2 + 3	717·1	—
5. Income	613·1	2
6. 4 − 5	104·0	—

This discrepancy is surprisingly large, but it is paralleled in almost all the income groups in which we have analysed our material, as is shown in Table 6, where the only group in which the discrepancy is of reasonable magnitude is the highest. We turn therefore to a consideration of possible reasons for the discrepancy.

TABLE 6

Mean Annual Expenditure, Saving and Income, by Income Group
(£ per household)*

	Income Group							
Item	1–3 £0–150	4 £151–260	5 £261–420	6 £421–620	7 £621–830	8 £831–1,040	9 £1,041–1,560	10 £1,561+
Food	67	98	150	192	228	259	295	304
Alcoholic drink	5	10	13	16	24	33	51	61
Tobacco	6	12	25	33	37	49	57	54
Housing	32	36	38	47	50	58	82	227
Fuel and light	22	23	29	32	36	39	46	71
Clothing	16	21	46	62	98	124	155	178
Travel	5	9	25	35	49	57	106	143
Major durable goods (net)	4	7	22	33	56	57	75	110
Other durable goods	4	7	22	25	37	48	60	101
Miscellaneous goods	11	15	31	30	38	51	55	64
Miscellaneous services	9	18	29	32	54	64	104	261
Other expenditure	21	32	41	58	88	96	157	328
1. Total expenditure	202	288	471	595	795	935	1,243	1,902
	(18)	(6)	(7)	(4)	(6)	(6)	(7)	(21)
2. Income tax and national insurance	4	5	20	31	53	87	170	516
3. Total saving	−25	−18	−11	7	21	39	114	40
	(24)	(32)	(47)	(77)	(42)	(33)	(43)	(360)
4. Sum of 1, 2 and 3	181	275	480	633	869	1,061	1,527	2,458
5. Total income	117	208	366	518	720	941	1,242	2,469
	(2)	(2)	(1)	(1)	(1)	(1)	(1)	(6)
4 − 5	70	70	121	121	156	127	296	10
4 as % of 5	155%	132%	131%	122%	121%	113%	123%	100%

* The figures in brackets are the percentage standard errors of the income group means.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Sources of Error

The first source of error to be considered is sampling error. Some of our individual items, even with a sample of over 3,000 households (about 1,500 each for income and saving) have quite considerable estimated sampling errors, but this is not the case for the totals of expenditure and income or for saving (where the proportionate error is large, but the absolute amount small). It is not possible for us to estimate the error of our £104 discrepancy, since it involves unknown covariances between expenditure and saving and expenditure and income (the estimates of income and saving are independent, being derived from different random samples). However, the covariance of expenditure and income must greatly exceed that of expenditure and saving, so the variance of the discrepancy can hardly exceed the sum of the variances of the individual items. On this assumption, a discrepancy of the size we have found is extremely unlikely to have arisen purely as the result of sampling variations either in the overall sample or in any of the income groups except the highest and, just possibly, the lowest (where the estimated error of expenditure is remarkably high).

A common source of error in sample surveys is non-response. We have a considerable number of refusals in our own survey which may affect the representativeness of our sample in various ways and bias some of our estimates. But the overall balance of expenditure, saving and income should apply to *any* group of households including the population of "co-operative households" which our sample represents. Non-response cannot therefore affect this particular discrepancy.

There are, however, two particular forms of non-response which occur in this survey and which could affect the overall balance. The first results from splitting the sample for our inquiries into income and saving. If the respondents in the two halves were very different types of household, this could upset our balance. This does not seem very likely, since people were not usually told anything about the questions which were to follow the expenditure details, so very few of the refusals can have known whether they were to be asked about savings or income. However, there were some differences in response, so we have compared the two halves of our sample in

TABLE 7

Household Income Distributions for Income and Savings Samples
(Weighted Population Estimate)

	Income Group							
	1–3	4	5	6	7	8	9	10
	£0–150	£151–260	£261–420	£421–620	£621–830	£831–1,040	£1,041–1,560	£1,561+
Income sample %	9·0	7·7	23·7	27·1	14·4	6·5	7·3	4·3
Savings sample %	9·9	9·4	21·5	26·2	16·0	7·1	6·9	3·0

TABLE 8

Mean Household Expenditure on Selected Items for Income and Savings Samples

Item	*Expenditure per Annum* (£)		*Std. Error (Whole Sample)* £
	Income Sample	*Savings Sample*	
Food	184·2	182·0	2·0
Life assurance	23·1	19·3	0·7
Major durable goods (gross)	44·4	39·7	1·8
Total outlay (including capital items) (week's estimate)	729·3	708·3	26·6

respect of some other characteristics. Table 7 shows the income distribution of the two halves by broad income groups, and Table 8 shows their mean expenditure on various items. The differences are small (a χ^2 test for the income distributions is not significant even when the groups

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

are amalgamated into three), but if a conclusion has to be forced out of them it can only be that incomes and expenditure are somewhat higher in the income half. If this were so, it would seem to follow that our estimates overstate income and understate saving, and therefore *understate* the discrepancy.

The second special form of non-response concerns those households which gave expenditure details and which are treated as co-operating but which refused (or were unable) to give sufficient details of income or savings to provide totals for those items. Total saving and total regular income are each not available for about 8 per cent. of their respective halves of the sample (7 per cent. in the lower rateable value strata and 10–15 per cent. in the higher). To estimate the bias which might result from this we have made a detailed study of these households for the income half of the sample. First we considered purely demographic characteristics, such as the average size of household and its age and sex composition. The general picture is one of smaller than average households except in stratum 3 (the low rateable value rural stratum) where they are larger than average; the excess in this stratum is among the 15–19 year olds and men aged 20–64, but the deficiencies in the other strata are all among children too young to earn. There is a high proportion of households consisting only of old people in the two high rateable value strata. Turning to socio-economic factors, we considered the proportion of households owning their own home or possessing a telephone, and the sex and employment status (employed, self-employed, unemployed or unoccupied) of the chief income receiver in the household (Tables 9, 10 and 11). The results for home-ownership are most striking. At first sight, it appears that the households for which we have no regular income totals must be in the highest income groups, but this is not borne out by the proportion possessing a telephone or by the fact that in strata 2 and 4 the proportion of chief income receivers who are unoccupied or women is markedly higher than average. There is, however, also a high proportion of self-employed chief income receivers among these households in all four strata, and this again is a feature which is associated with the higher income groups. We have concluded from this that these households, for which we do not have an income total, tend to be clustered towards the two ends of the income scale. This does not conflict with the home-ownership figures—since the proportion of home-owners is also fairly high at the bottom end of the income scale—and it seems also to be borne out by the size distributions of a number of expenditure items which we have examined. This conclusion implies that, although these households have a number of special features, their *average* income cannot be very different from that of the rest of the sample; they cannot, therefore, have much bearing on the discrepancy we are considering.

TABLE 9

Proportion (%) of Households Owning Home

	Income Group									
Stratum	1–3	4	5	6	7	8	9	10	*Whole Sample*	*Households Without Regular Income Total*
1	34·0	31·7	27·4	28·0	36·3	42·2	43·0	42·9	33·1	47·7
2	—	14·3	7·7	12·1	28·0	24·0	42·1	64·5	33·9	55·6
3	34·8	42·0	17·9	22·3	29·4	35·7	39·0	55·6	28·4	43·9
4	34·8	32·0	23·5	35·1	51·1	36·0	58·5	75·0	45·7	68·2

TABLE 10

Proportion (%) of Households with Telephone

	Income Group									
Stratum	1–3	4	5	6	7	8	9	10	*Whole Sample*	*Households Without Regular Income Total*
1	3·1	7·7	10·0	9·7	17·0	25·6	44·3	71·4	15·2	20·5
2	33·3	14·3	46·2	34·8	46·0	60·0	66·7	93·4	59·2	66·7
3	3·2	8·9	7·0	9·6	16·2	25·7	30·5	77·8	11·9	31·7
4	21·7	22·2	23·5	43·2	61·7	68·0	81·1	97·5	55·6	59·1

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 11

Employment Status of Chief Income Receiver

(Proportion (%) of Chief Income Receivers of each Employment Status in each Income Group)

Stratum	*Employment Status**	*Household Income Group* 1–3	4	5	6	7	8	9	10	*Whole Sample*	*Households Without Regular Income Total*
1	1	11·3	35·6	78·5	90·6	91·5	94·4	86·1	71·4	76·4	75·0
	2	2·1	5·8	6·7	5·1	5·4	2·2	12·7	21·4	6·4	18·2
	3	1·0	2·9	—	—	—	—	—	—	0·3	—
	4	85·6	55·8	14·8	4·3	3·1	3·3	1·3	7·1	16·9	6·8
2	1	—	28·6	69·2	69·7	72·0	80·0	70·2	82·9	69·1	33·3
	2	16·7	28·6	15·4	15·2	20·0	8·0	15·8	3·9	14·1	38·9
	3	16·7	—	—	—	—	—	—	—	0·3	—
	4	66·7	42·9	15·4	15·2	8·0	12·0	14·0	13·2	16·5	27·8
3	1	8·4	29·5	84·7	83·8	80·9	81·4	74·6	44·4	66·1	39·0
	2	1·3	8·9	10·2	13·4	17·6	15·7	25·4	55·6	13·5	51·2
	3	0·6	2·7	0·3	—	—	—	—	—	0·5	—
	4	89·7	58·9	4·8	2·7	1·5	2·9	—	—	19·9	9·8
4	1	43·5	55·6	70·6	51·4	57·4	72·0	54·7	40·0	54·0	27·3
	2	4·3	3·7	21·6	27·0	27·7	24·0	34·0	37·5	26·9	59·1
	3	—	—	—	—	—	—	—	—	0·3	—
	4	52·2	40·7	7·8	21·6	14·9	4·0	11·3	22·5	18·8	13·6

* 1 = employed; 2 = self-employed; 3 = unemployed; 4 = unoccupied.

We have not made a similarly detailed study of the households which have not given us total savings figures. It is clear, in view of the known order of magnitude of household saving, that any bias from this cause could have only a small influence on the discrepancy.

Non-Statistical Biases

If the discrepancy cannot be explained in terms of these sample biases we must consider non-statistical sources of bias. At a time of rising prices the use of a continuing sample over a year, with estimates obtained from data relating to past periods of varying length, may well give rise to some discrepancy. Although all three of our major aggregates are obtained from different combinations of past periods, downward bias is likely to be more important in the case of income and saving than in that of expenditure (as can be seen from the list of periods in Table 1). However, much of our income information is at current rates, and it seems unlikely that this particular source of bias has led to understatement of total income by much more than 1 per cent. The corresponding downward bias in saving (where all our information relates to the past) will tend to offset this, so the net effect on our "gap" will be small.

The most important aspects of reporting errors, memory omission and end-period effect, have already been touched on. Memory omission is straightforward. The respondent fails to recall the fact or the amount of expenditure, income received, etc., in the survey period. It leads to downward bias in the recorded answers. All commonly used survey methods rely to some extent on memory, but there is a general presumption that the shorter the memory period the less the danger of downward bias—hence the preference for record keeping where, in theory, the memory period is a day.

The end-period effect may also involve memory in the sense that the respondent genuinely telescopes time and believes that certain activities relate to the survey period when they really belong to an earlier period. But it also involves psychological factors—a desire on the part of the respondent to impress, an excessive desire to be helpful and to provide as many positive answers as possible. These features help to explain why even record keeping, where the beginning and end of the period are strictly delimited, still appears to be subject to end-period effect in at least the first week. Whatever the cause, the result of end-period effect is upward bias.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Another source of bias with any survey method is the reluctance of the respondents to reveal information about particular items. This inevitably leads to downward bias and well-known examples of such items are expenditure on alcoholic drink and tobacco, and income in general, but particularly certain parts of income—overtime earnings, rent, dividends and interest. Sometimes, however, this reluctance to reveal information cannot be distinguished from the fourth source of bias—namely the general difficulty for respondents of understanding the concepts involved. The concept of "business saving" for the self-employed is a case in point. Here, however, it is impossible to generalize about the direction in which the bias will operate.

The nature of these reporting biases is familiar enough. What we have no idea about is their relative magnitude or the *net* effect of biases operating at the same time but in different directions. We shall consider whether the S.A.C. inquiry can cast any light on these questions and whether they can help to explain our gap. We shall first compare our results with the few external estimates which are available.

External Comparisons

There are no figures relating to the administrative county of Cambridgeshire with which we can compare our results. Comparisons with national aggregates cannot be precise because regional variation is always possible and, indeed, likely. But it is unlikely to exceed certain limits, and local knowledge will help us to say for what items and in which directions it is most likely to occur.' We shall therefore make a comparison between the S.A.C. results and the Central Statistical Office (1955) estimates of personal expenditure, income and saving in the United Kingdom for the calendar year 1953, the National Food Survey for the calendar year 1953 (Ministry of Agriculture, Fisheries and Food, 1955) and the Oxford Institute of Statistics Savings Survey for the financial year 1953/54 (Straw, 1955, p. 288; Hill, 1955, p. 138).

(a) *Expenditure Estimates*

For this purpose the C.S.O. figures have been converted to a *per capita* basis and certain adjustments made to the commodity groups both for the C.S.O. and S.A.C. to make them as comparable as possible. Even so, certain differences remain which can be ignored for such a general comparison as we intend to make. The results are shown in Table 12.

The Cambridgeshire estimates of total expenditure are markedly higher than the adjusted C.S.O. total and there are considerable differences between commodity groups. Expenditure on food, alcohol and tobacco are all lower in the S.A.C. estimates, whereas for clothing, durable goods and the miscellaneous group they are higher.

We have noted the general difficulty of estimating expenditure on alcohol and tobacco in surveys of this kind. On the other hand, there is always a tendency to regard the C.S.O. figures for these commodities as accurate because they are derived from information about releases from bond. To arrive at consumers' expenditure, however, an allowance has to be made for business consumption which may possibly be under-estimated. Even this, together with an allowance for some regional variation in consumption, would hardly account for the lowness of the S.A.C. figures, and we must conclude that the answers to these questions are biased downwards. Respondents do not usually have the same reluctance to reveal food expenditure. The C.S.O. figures include an allowance for home-produced and consumed food which S.A.C. does not include and which will clearly be important in an agricultural county. But the S.A.C. figure is also some 6·0 per cent. lower than the National Food Survey's estimates of *purchases* of food in 1953, the main deficiencies being on fruit, meat and bread.

On practically every other commodity group S.A.C. expenditure is higher than the national estimates. This is particularly marked in the case of the miscellaneous group where we are unable to make any more detailed comparison. The Cambridgeshire figures here include such things as builders' charges for interior decoration (which are in the C.S.O. figures for housing), hotel expenses, bank charges, legal fees, toilet and chemists' goods, etc. It is quite likely that the C.S.O. estimates will not be particularly reliable for such a group, where information is hard to come by, but it is difficult to believe that they cover only 50 per cent. of such outlay, or alternatively, that Cambridgeshire households spend so very much more than the average.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 12

Comparison of Cambridgeshire and National Estimates of Consumers' Expenditure

Item	*C.S.O.* 1953 *Annual* per Capita £	*S.A.C.* 1953/54 "*Last week*" *Annual* per Capita £	*S.A.C.* 1953/54 "*Period*" *Annual* per Capita £
Housing	19·6	18·3	19·3
Fuel and light:	8·9	15·3	11·7
Gas	2·1	2·9	2·8
Electricity	2·2	3·9	3·1
Coal, coke and other fuel	4·6	8·5	5·9
Food	70·3	64·6	—
Alcohol	16·4	7·3	—
Tobacco	16·5	11·0	—
Books and newspapers	3·5	3·5	—
Travel:	12·2	13·6	14·0
Vehicle running costs	3·6	—	9·3
Fares	8·6	—	4·7
Entertainment	3·3	5·3	—
Postage, telephone	1·5	2·0	1·9
Soaps, polishes, etc.	2·7	4·6	—
Domestic service	1·9	3·0	3·4
Clothing	21·1	19·4	25·4
Major and other durable goods	20·0	39·9	25·0
Miscellaneous goods and services, by difference*	16·8	27·8	30·2
Total	215·7	235·6	227·2

* Excluding fines, pocket money, etc.

It is generally true that, even making allowance for some error in the C.S.O. estimates, the S.A.C. period estimates of a number of the commodity groups look much more reasonable than the week's estimates. One should bear in mind that £1 *per capita* is £50 million nationally, so that the S.A.C. week's estimate of major durables would represent a national total of some £2,000 million—when at the moment the 1953 estimate of total consumers' expenditure on all items amounts to only just over £11,000 million. Although such a comparison is inevitably inconclusive, it confirms our view that of the two S.A.C. estimates the period ones are to be preferred but that even these are high.

(b) *Income Estimates*

The following table shows the results of a comparison between the S.A.C. income figures, C.S.O. estimates and the Oxford Savings Survey for 1953/54 all on a *per capita* basis.

TABLE 13

	S.A.C. 1953/54		*Adjusted C.S.O.* 1953		*Oxford Savings Survey* 1953/54	
	£	*Percentage of Total Income*	£	*Percentage of Total Income*	£	*Percentage of Total Income*
Wages and salaries	158	72·8	168	67·2	161	77·
Self-employment income	22	10·1	29	11·6	21	10·1
Rent, dividends, interest	14	6·5	28	11·2	10	4·8
Government transfers	15	6·9	17	6·8	15	7·2
Other income	8	3·7	8	3·2	—	—
Total	217	100·0	250	100·0	207	100·0

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

One or two points of definition should be noted. The Oxford Survey figures are estimated totals for Great Britain, reduced to a *per capita* basis by dividing by the population of Great Britain at mid-1953. The C.S.O. figures are obtained, as were the expenditure figures, from United Kingdom totals. The Oxford figure for wages and salaries includes private pensions, while these are shown separately (as the sole constituent of "other" income) by the C.S.O. In the Cambridgeshire figures private pensions are included in "other" income, along with regular transfers of income from relatives and friends outside the household (excluded altogether from the C.S.O. figures and from the Oxford estimates used here) and gross income from lodgers and boarders living in the same house. Lastly, the Oxford estimates, being derived from income units living in private households, exclude the income of the institutional population, although for reducing them to a *per capita* basis they have been divided by the total population. Oxford's own estimates suggest that the total of £207 per head should be increased by about £7 to allow for this, although it is impossible to say how this is spread over the various types of income.

Bearing these factors in mind it appears that the Cambridgeshire estimates are quite close to the Oxford estimates both in total and in the component parts. When compared with the C.S.O. figures, however, they are very low. The major discrepancy is in the rent, dividends and interest figure, which is only 50 per cent. of the corresponding C.S.O. figure, while self-employed income and wages and salaries are some 25 per cent. and some 6 per cent. lower. The S.A.C. total is 7·5 per cent. lower than the corresponding C.S.O. total, whereas data derived from the Inland Revenue 1949–50 income census (Commissioner of H.M. Inland Revenue, 1953) show a Cambridgeshire average income (and here Cambridgeshire includes the Isle of Ely) some 4 per cent. lower than the United Kingdom average.

Most of the evidence points to wage and salary incomes in Cambridgeshire being rather lower than the national average. Agriculture is very important in the economy of the county. Some 13 per cent. of the occupied population in our sample was engaged in agriculture compared with 4·7 per cent. of the occupied population of England and Wales in 1951. In 1953 the national minimum agricultural wage was £6 10*s.* a week for males compared with an average £9 9*s.* for all industries. Other wages in Cambridgeshire also appear to be on the low side, and there is a higher than average proportion of old people, which may well mean more old age pensioner households with low incomes. Against this must be set the fact that there is a concentration in the county of professional workers, particularly University lecturers, etc., and of civil servants. But the salaries of University staff are rather low compared with other similar professions, and the civil servants are almost entirely executive and clerical. Whether regional variations are sufficient to account for all the observed differences between our figures and those of the C.S.O. it is impossible to say.

Self-employment income might have been expected to play a more important role in total income in Cambridgeshire than in the country as a whole because there are relatively more self-employed, and there is not much evidence that Cambridgeshire self-employed incomes are lower than average. In the case of rent, dividends and interest there is virtually no evidence for regional variations other than that suggested by the absence of really large incomes which are likely to contain a high proportion of unearned income. It would be difficult to account for the whole difference in this way and it seems much more likely that the Cambridgeshire survey, like most income surveys (British and American) has failed to obtain reliable estimates of income from these sources.

(c) *Saving Estimates*

Detailed comparison of the savings estimates of the S.A.C. inquiry with C.S.O. figures is not very instructive. In the first place the C.S.O. readily admits the weakness of its own estimates (Saunders, 1954) and, although some components of saving are now identifiable and estimated directly, the estimate of total personal savings is the difference between the estimates of personal income and of personal expenditure. Secondly, practically nothing is known about regional variations (if any) in savings patterns. However, the 1953/54 Oxford savings survey covers the same period as our own, and its methods are so similar that comparison with its results may be more informative.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 14

	S.A.C. 53/54 £ per capita	*Oxford* 53/54 *National* £ per capita
Total saving	+3·4	+5·2
Contractual saving	+8·6	+9·9
Change in liquid assets	−6·9	−5·2
Other saving	+0·7	−0·9

The similarity of these results is confirmed by a comparison of the relationship between gross income and saving for the sample units in the two surveys (see Diagram I).* Both Oxford and Cambridge estimates of total saving look low in relation to the C.S.O. figure of £12 *per capita*,

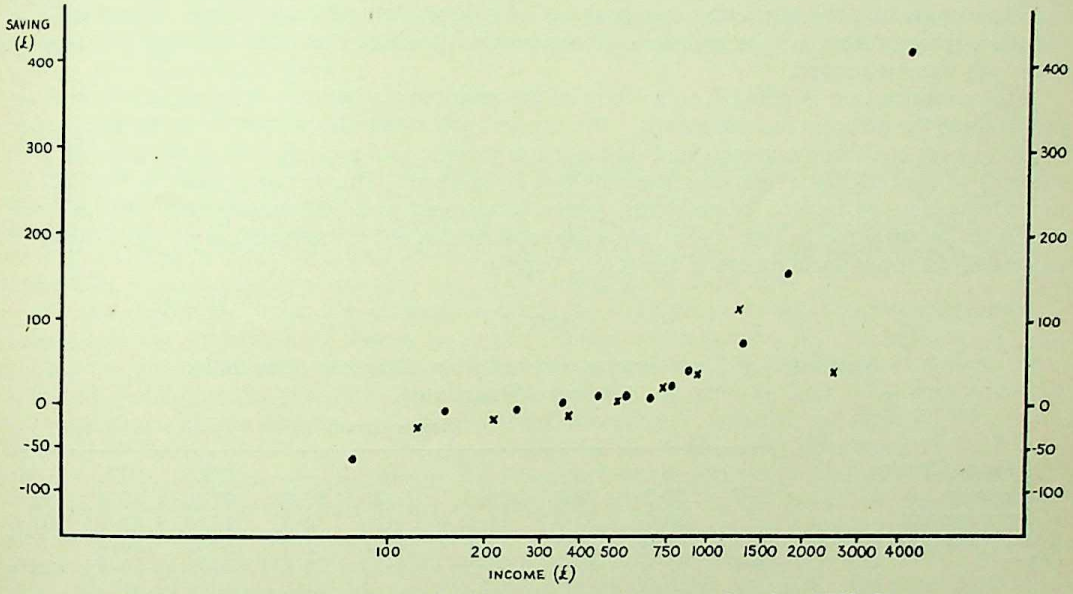


DIAGRAM 1. — The relationship between gross income and saving, 1953–54.
● Oxford savings survey.
× S.A.C. inquiry.

but this possible under-estimation of saving in the S.A.C. figures is negligible in relation to the gap and operates in the wrong direction.

We may conclude from these rough comparisons that all the blame for the lack of balance cannot be laid at the door of income. Income is under-estimated but expenditure is also over-estimated. To put the very roughest order of magnitude upon the discrepancies we might say, but only as a guess, that the S.A.C. total expenditure needs to be reduced by 5 per cent. and total income increased by some 10 per cent.

Biases in the Components and between Income Groups

The comparison also makes clear that the incidence of bias between the components of income and expenditure is uneven. In fact, if we allow for the obvious under-estimation of expenditure on tobacco and alcohol the estimates for the remaining commodity groups need to be reduced by about 10 per cent. Among these there are some, such as housing and food, which seem fairly reasonable. This means that the remaining items accounting for only 55 per cent. of the S.A.C.

*We are indebted to Mr. L. Klein for permission to use these unpublished data from the Oxford Savings Survey.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

total expenditure have to bear the full weight of over-estimation. Among these items the over-estimation also appears to operate unevenly, so that expenditure on some particular commodity groups may need to be reduced by 20 per cent. or more.

It may be that the psychological and sociological aspects of end-period effect influence some particular commodities more than others. In budget surveys among peoples with a low standard of living end-period effect appears to operate over practically all expenditure items. The experiment of Mahalanobis and Sen (1954, p. 9) revealed that in Indian conditions a month's reference period was to be preferred to a week for certain food items. Nicholson (1954, p. 71) in describing the results of his work among Arab families in Tripolitania suggested that all commodity groups were affected by over-statement. In a country like the United Kingdom, however, it is clearly the items like durable goods which are particularly memorable or are objects with which people wish to impress. At the same time, it seems very probable that there will be differences between income groups in this respect. Certainly one would expect the desire to impress to be less in the higher income groups and the particular memorability of a large item of expenditure, which leads to the telescoping of time, will be greater for those people for whom it is more unusual, i.e. those in the lower income groups.

This consideration is relevant to a study of the relationship between income, saving and expenditure in the different income groups. We saw in Table 6 that the proportionate gap tended to diminish over the income groups and that our top income group consisting of households with incomes of over £1,560 a year was the only one in balance. This is unexpected in the light of the comparison of income components, where investment and self-employment income were found to be particularly low. The following table shows the concentration of these types of income in the top income groups of the S.A.C. inquiry.

TABLE 15

Percentage of Total Income Derived from Different Sources by Income Group

	Income Group								
Type of Income	1–3	4	5	6	7	8	9	10	All
Employment . . .	5·7	24·2	76·2	79·3	81·8	87·1	72·9	57·4	66·6
Self-employment . .	—	10·8	5·9	6·2	7·7	4·6	12·0	19·1	10·5
Rent dividend and interest	10·2	5·4	1·7	2·6	2·2	2·7	7·6	20·4	6·5
Other	84·1	59·6	16·2	11·9	8·3	5·6	7·5	3·1	16·4

The 94th Report of the Commissioners of Inland Revenue (1952) enables us to estimate the proportion of income which was derived from different sources in 1949–50. Investment income accounted for 34 per cent. and self-employment for 32 per cent. of the total income of tax units with incomes over £1,500, compared with S.A.C estimates of 20 per cent. and 19 per cent. for households in the range £1,561 +.

This suggests that income from rent, dividends and interest and from self-employment is in fact considerably understated in our top income group, and that this will account for a good deal of the total understatement of income from these sources. To go beyond this and estimate the extent of understatement in this and other income groups we have to make some very arbitrary assumptions. We might assume that the understatement of investment and self-employment incomes is *proportionally* the same in all income groups, and of the order of magnitude indicated by the comparisons in Table 13, and that income from employment and government transfers is not understated at all in the top income group but is somewhat understated (4 or 5 per cent.) in the other groups. In round terms these assumptions suggest that the S.A.C. estimates of total income should be increased by some 10 per cent. in each income group except group 5 (£261–420), where the deficiency is somewhat less, and group 10 (over £1,560) where the increase should be as much as 20 per cent.

Such an adjustment to income has interesting implications. In the first place the balance of income group 10 disappears and is replaced by a large excess of income over expenditure and saving. Some of this can almost certainly be accounted for by the low estimate of savings for this group. We saw in diagram I that the relationships between income and gross saving for the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Oxford and Cambridge data are very similar in every income group except the highest where, in the Oxford figures, savings continue to rise whereas the S.A.C. estimates show a fall. Subsequent work of the Oxford savings survey confirms the heavy concentration of savings among the very wealthy. Thus it seems that the S.A.C. savings figure should be considerably increased; but at the same time some under-estimation of expenditure in this income group should not be ruled out. In the other income groups the estimate of savings may be somewhat too high, but this cannot greatly affect the discrepancy, and expenditure is almost certainly considerably over-estimated. For instance, if our estimate of the dis-saving of the lowest income group is correct, their expenditure would have to be reduced by something like 25 per cent. to arrive at a balance with the adjusted estimate of income. A similar situation prevails in the other income groups, although the proportion by which expenditure has to be reduced declines fairly steadily over the income groups until the probable *under-estimate* is reached in the top income group. In view of the considerable uncertainty attaching to the savings figures we cannot be more precise than this; but these conclusions certainly lend support to the *a priori* reasoning which led us to expect end-period effect to operate more generally and to have more influence on the estimates of expenditure for the lower income groups.

Conclusions

The phenomenon considered in this paper is sufficiently widespread for it to be clear that the weaknesses revealed in the S.A.C. data are by no means confined to this inquiry. We have deliberately set out to emphasise these limitations in order to focus attention on the problem of improving the data in the future, and to draw the attention of users of existing budgetary data to some of the main biases for which they should be making some allowance.

Hitherto in family budget studies most attention has been paid to ways of increasing response. Record keeping methods have secured such poor co-operation that in the United States at any rate there is now a general preference for the interview method. The S.A.C. inquiry was the first large-scale family budget inquiry in Britain to use only the interview method. The response rate of 75 per cent. was encouraging. For purposes of comparison, however, we must note that the Ministry of Labour 1953 household budget inquiry in which primarily the record method was used achieved a response of 67 per cent. At the same time the expenditure information collected seems to have been, for the most part, reasonably free of bias. In the *Report on Proposals for a New Index of Retail Prices* (Ministry of Labour and National Service, 1956) it is stated that apart from some understatement of expenditure on tobacco, alcohol, meals eaten out, sweets and ice cream, the recorded expenditure agreed closely with other available statistics. On the other hand, income appears to have been understated. Response bias and reporting bias are not, of course, the only considerations. Cost is also a factor of some importance, and presumably the high Ministry of Labour record-keeping response is not independent of the payment that was made to respondents. But a comparison of this kind does raise the question of which method is to be preferred.

A natural reaction, in view of the discrepancies we have discussed, might be that collection by any method is so unreliable that family budget studies are useless. It is probably true that for the estimation of many national aggregates other methods are superior. But for regional or income group comparisons, to name but two possible uses, family budgets are indispensable. The S.A.C. inquiry would suggest that unless *very* marked income differences are to be expected the present methods might be quite useful for comparisons between regions, but for income group comparisons the data can only be used with serious reservations.

The main conclusion seems to be the need to devote further attention to the elimination (or estimation) of these reporting biases, and particularly the end-period effect. We feel that it would be wrong at present to write off the interview method in favour of record-keeping. The latter has usually achieved a lower response rate—neither is really satisfactory—and is certainly subject to some of the same biases, particularly in respect of income. In our own survey it seems clear that the reporting biases are more important than sampling error, although they may themselves affect our estimates of the latter. If this is already true in a sample of 3,000 households, it is clearly important that in future surveys of this kind resources should be devoted to the reduction of reporting biases rather than to an increase in sample size.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Acknowledgements

We are indebted to a number of people who have helped in various ways with the very large body of tabulations and computations involved in the compilation of the estimates given in this paper. We should like particularly to thank Mr. J. A. C. Brown, who evolved a programme for the electronic computation of our means and sampling errors, and gave much time to carrying out this work for us, and also Mr. J. L. Nicholson for his constructive criticism of the original draft of this paper.

References

Central Statistical Office (1955), *National Income and Expenditure*. London: H.M.S.O.
Commissioners of H.M. Inland Revenue (1952), 94*th Report*, Cmd. 8436. London: H.M.S.O.
—— —— —— (1953), 95*th Report*, Cmd. 8726. London: H.M.S.O.
EDEN, Sir F. M. (ed. Rodgers, 1928), *The State of the Poor*, abridged. ed. London: Routledge.
HILL, T. P. (1955), "Incomes, savings and net worth: the savings surveys of 1952–54", *Bull. Oxford Inst. Stats.*, **17**, 129–172.
MAHALANOBIS, P. C. and SEN, S. B. (1954), "On some aspects of the Indian National Sample Survey", *Bull. Int. Statist. Inst.* **34** (2), 5–14.
MASSEY, P. (1942), "The expenditure of 1,360 middle-class households in 1938–39", *J.R. Statist. Soc.* A., **105**, 159–185.
Ministry of Agriculture, Fisheries and Food (1955), *Domestic Food Consumption and Expenditure*, 1953. London: H.M.S.O.
Ministry of Labour and National Service (1956), *Report on Proposals for a New Index of Retail Prices*, Cmd. 9710. London: H.M.S.O
NICHOLSON, J. L. (1954), "A survey of the living conditions of Arab families in Tripolitania in Nov.–Dec., 1950", *Rev. Int. Statist. Inst.*, **22**, 68–84.
PRAIS, S. J. and HOUTHAKKER, H. S. (1955), *The Analysis of Family Budgets*. Cambridge: Cambridge University Press.
SAUNDERS, C. T. (1954), "Some problems in the estimation of personal savings and investment", *Trans. Manch. Statist. Soc.*, 54–55.
STRAW, K. H. (1955), "The savings surveys and official sources: a reconciliation of estimates of personal income", *Bull. Oxford Inst. Stats.*, **17**, 283–302.
U.S. Bureau of Labor Statistics (1952), "Survey of consumers' expenditure in 1950: interpretation and use of the results", *Monthly Labor Review*, **75**, 425–428.
UTTING, J. E. G., and COLE, DOROTHY (1954), "Sampling for social accounts—some aspects of the Cambridgeshire survey", *Bull. Int. Statist. Inst.*, **34** (2), 301–328.

APPENDIX

The Main Contents of the S.A.C. Commodity Groups

Housing expenditure	Rent, rates, water rates, external repairs, mortgage interest, house insurance.
Food	All usual food items and meals eaten out.
Alcoholic drink	Beer, wines, spirits.
Tobacco	Cigarettes, tobacco, snuff.
Fuel	Gas, electricity, coal, coke, paraffin.
Soaps, polishes, etc.	All cleaning materials, matches, electric light bulbs, etc.
Books, newspapers	Newspapers, periodicals, books, book tokens.
Postage and telephone	—
Miscellaneous services	Builders' charges for interior decoration, legal charges, interest other than mortgage, repair and maintenance of everything except vehicles and houses, hotel, removal and funeral charges, hire charges, hairdressing, laundry and dry cleaning.
Miscellaneous goods	Paint, distemper, toilet accessories, smokers' requisites, stationery, seeds, flowers, animal foods.
Entertainment	Cinema, theatre, entrance to sporting events, gross betting expenditure, football pools.
Transport	Railway and bus fares, taxis, maintenance and running of vehicles including road fund licence and insurance.
Clothing	All clothing, haberdashery, knitting wool, material, etc.
Money gifts, transfers, etc.	Gifts less than £5, subscriptions to non-profit-making bodies and to charities.
Domestic service	Regular and casual, including window cleaning, baby sitting.
Property insurance	Other than vehicles and houses.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Durable goods:	
Major durables . . .	Refrigerator, washing machine, vacuum cleaner, cooking stove, boiler, car, cycle, pram, boat, caravan, television, radio, piano, typewriter, sewing machine, furniture, carpet, linoleum, lawn-mower, garden shed, garage.
Other durables . . .	Optical goods, toys, sports equipment, leather goods, jewellery, household linen and soft furnishings, pots and pans, gas fires, electric fires, clocks, kettles, etc., garden tools.

Discussion on Paper by Mrs. Cole and Mr. Utting

Professor Allen: It is my privilege to be the first to welcome this paper by Mrs. Cole and Mr. Utting. Household budget inquiries are a field in which I have done some work and the particular topic discussed, expenditure plus saving in relation to income, is a nice blend of the methodological and the practical. It is also good to have a discussion on survey methods which is not completely bedevilled by the question of non-response, and equally good to have a paper which is short, at least relative to some of the papers presented to this Society. The problem here is to recognise and to assess discrepancies in budget survey data with the object of eliminating or, more usually, of allowing for them. The authors throw a good deal of light on the relative merits of the interview as opposed to the record-keeping method. Record-keeping in budget surveys may be out of fashion in the United States but I would suggest that it is probably best to proceed by some combination of the two methods, as adopted, for example, by the Ministry of Labour in their 1953 Inquiry.

The question of timing of expenditures is very important in the interview method when an aggregate is made up of seperate estimates spread over weeks, months, quarters or years. If prices and incomes are rising then discrepancies in the relation of expenditure plus saving to income can arise for this reason alone. I was grateful to Mrs. Cole for adding some comments on this particular point and I hope that these will appear in the paper when it is finally published.

I do not think the authors would claim to have had foreknowledge but at least they were extremely fortunate in the choice of their years. The period from 1952 to early 1954 was the last period of relative stability in prices and incomes in this country—not the last for all time, I hope—and the effect of timing will not be at all important in their results. The Ministry of Labour were equally fortunate in the choice of the same year for their household budget inquiry.

It is probably inevitable that brevity implies an occasional statement which is unsupported or even dogmatic and such, I think, are some of the authors' comments on the end-period effect. It seems to me that they take it too much for granted that the evidence is conclusive, that end-period effects will arise even in record-keeping, causing some items of expenditure to be over-estimated by very significant amounts. If the record-keeping method is adopted, and if the design of the inquiry is a good one, I am not at all convinced that the end-period effects will be serious. I am speaking here with some knowledge of the Ministry of Labour's 1953 inquiry and it will be interesting to see what conclusions the Ministry come to on this point when their analysis of the 1953 budget data is completed.

If I may conclude with a parenthesis, I note the authors' statement about low university salaries; that is one dogmatic statement I am quite prepared to accept without further proof.

It is with very great pleasure that I move a vote of thanks to Mrs. Cole and Mr. Utting for presenting this interesting paper.

Mr. Saunders: It is a pleasure to second this vote of thanks to Mrs. Cole and Mr. Utting. It seems to me that the present stage of household budget inquiries calls for exactly this kind of paper. It records an investigation, not so much to give substantial results, as to try and identify what sort of social and economic information can be obtained by this method. The paper must be read in conjunction with the other papers which have been produced resulting from the Cambridgeshire inquiry.

I think it has to be recognized—the authors have recognized this—that with present techniques household budget inquiries are really rather a blunt instrument. Their value for broad analyses of household expenditure cannot be doubted any more, apart from the well-known difficulties of getting honest statements about expenditure on drink, tobacco and betting; but when we get to the delicate subject of comparing income with expenditure, the paper shows that the technique has reached a barrier which is perhaps technical but must be very largely psychological. One gets exactly the same impression from the Oxford Institute of Statistics' Savings Inquiry and also from the corresponding Michigan Surveys in the United States. Yet in spite of this the possibility

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of using budgets for refined and delicate analyses of saving and for calculating saving functions, and ultimately, one would hope, for calculating national aggregates, is an extremely attractive one. It is a possibility which one cannot dismiss as being beyond the scope of sample surveys, but it does need some very substantial advances in technique.

Several questions arise. The present paper uses a double method; some people have been asked by the Survey for their expenditure and their savings, from which income is deduced; some have been asked for expenditure and income, from which savings is deduced. The Oxford survey asked for savings and income, from which expenditure can be deduced.

Can we take the final step of venturing to ask each household in a sample for all three items—income, expenditure and saving—that is, asking respondents to produce a current account and a capital account? The advantage of this is not only the completeness of the information but that the skilful interviewer has the opportunity of getting the respondent to make sense of the answers. Is there anybody here, fortunate enough to have any assets or liabilities, who has tried to reconcile his own current account with his capital account at the end of the year? From personal experience, I would say that what goes wrong every time is under-estimation of expenditure.

The Oxford Survey shows clearly that if one is interested in assets and saving, the problem comes down to the people with £1,000 a year and more, since that is where the saving is concentrated. It would not be easy to get the majority of people who earn £1,000 a year to produce reconciled current and capital accounts. But unless one can somehow get reliable data from people who have a really complex structure of assets and liabilities, the pursuit of a reliable figure of savings by sample surveys is bound to fail.

I am not entirely happy that the question of the differential rates of non-response as between income groups has been fully dealt with. It is quite true that it does not affect the existence of the discrepancy between income and expenditure, as the authors say, but there is evidence that the discrepancy is greater among the higher income groups; small differences can substantially affect the level of saving.

Great difficulty must arise in dealing with self-employment income. There is no attempt, as there was in the Oxford Survey, to separate the self-employer as a business man from the self-employer as a contributor to personal saving. In the Cambridgeshire Survey, what is put back into the business is treated as a form of increase in assets which seems the right way to deal with it.

The great difficulty of getting up-to-date figures from self-employers—of whom many do not seem to be able to produce up-to-date business accounts—makes one distrustful of the resulting income figures. Possibly a special approach is needed to small businesses, with the object of reconstructing the whole trading account and deriving from it a sensible figure of profit. To ask directly for a figure of income is open to a great deal of ambiguity.

The paper is very stimulating and it will be studied for a long time. It gives a great many hints about the breaking of the barriers in these inquiries.

The vote of thanks was put to the meeting and carried unanimously.

Mr. W. F. F. Kemsley: One reason for welcoming this paper is that it emphasizes the importance of reporting errors instead of concentrating on sampling. Errors occur in reporting in budgetary and allied surveys. This can be demonstrated by carrying out suitable experiments. The real problem is how do errors occur and how can they be overcome. The inclusion of savings is an improvement on earlier investigations but it is not necessary to bring in savings to show that something is wrong.

If we turn to Table 6 and take the total of income tax, National Insurance and total expenditure, and subtract from total income, we get, reading from left to right:

89, 85, 125, 108, 128, 81, 171 and − 51

These do not appear to be associated with income but to vary randomly about a constant of £90. I suggest that there is a simple explanation, that households exaggerated their expenditure by £2 a week irrespective of the size of their total budget and forgot that savings must be met out of current income and they did not reduce their expenditure as they should have done. One may ask whether some of the bias could have been avoided if each household had been asked to give a complete account of savings, expenditure and income. Presumably this was not done because it might have brought about a lower response, but even if a balance had been attempted it does not follow that the data would have been more acceptable. In this connexion some account of the experience of the Social Survey some years ago may be of interest.

During the course of experimental budget surveys in 1950–51 we tried a reconciliation of this kind. We found that in any representative group of households expenditure on the average was

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

10 to 20 per cent. above income. The figures on p. 376 give a corresponding percentage of about 15 per cent. The reconciliation was limited to cash because we were anxious not to depress the response which we feared might have happened if all transactions through bank accounts were inquired into. The cash reconciliation was made as complete as possible, including withdrawals in cash from savings together with cash in hand at the beginning and end of the week. A balance was drawn up for each individual separately taking account of cash transfers within the household. Although this was not a complete check, we hoped to remove some at least of the discrepancies by reducing the inflation of expenditure and possibly picking up undisclosed income. It was not a success. There was an excess of cash outgoings over cash receipts of the order of 6 per cent. when measured in terms of household income. The failure was explained by the difficulty of drawing up a balance sheet while interviewing. It was difficult to get a correct total of the items of expenditure, and as a result some discrepancies were not detected and others were found where none existed. The additions were eventually made in the office but we were not able to call back to re-check or repeat the experiment. Re-calling for the purpose of checking discrepancies has certain disadvantages. The delay will increase the risk of defective memory and the extra calls make it more expensive and possibly increase resistance.

I intended commenting on the end-period effect of memory. I was struck with the figures in Table 2 on clothing. Mrs. Cole has already referred to them, pointing out that the last week's estimate of clothing is less than the period estimate. In various checks that we have done from time to time it seems to be a general rule that the level of expenditure is reduced by increasing the memory period. The difference in the table is larger than can be attributed to sampling errors and presumably the last week's estimate has been affected by the variability of clothing expenditure. I gather that quarterly variations have been balanced by giving equal weight to each quarter but presumably the effect of variations within the quarter still remains because the field work will not have yielded the same number of households for each week. I wonder whether the authors can tell us how far this would explain the lower estimate.

Miss M. DEANE: I should like to congratulate Mrs. Cole and Mr. Utting on the practical aspect of their paper. I should have liked more details on how they collected the income side of it. I am not quite satisfied that there is not a time lag here of the kind which does not get discounted by the fact mentioned by Professor Allen—that this period was a time of stable prices. For instance, if a retired schoolmaster on a low pension declared a dividend, was full account taken of the fact that the full rate of income tax had been deducted at source, and that such a person would eventually get a rebate? I should have thought that it would be extremely difficult to get a real tie-up in many small ways such as this. I should also like to ask whether any breakdown of the expenditure has been done according to the size of family or to the employment status of the head of the household? Both would seem to me quite likely to show some sort of correlation.

It seems to me—I have no experience of surveys, I am merely a potential user of these statistics—that information on seven days' expenditure by interviewing does not give a real picture, that the big discrepancy apparent in some of the period estimates and weekly estimates is likely to occur in other items. I confess that if somebody asked me to give an account of my expenditure for the last seven days I should not remember whether it was seven or eight days since I bought an article such as face powder. It looks as if some combination of interview and record method or some check would have to be made.

Although perhaps at the moment it all seems full of small snags and rather discouraging, I am sure that there is a big demand for these regional budgets. I am frequently asked whether I can supply sources of information of regional differences in the pattern of spending or in the cost of living and have to say "No", but perhaps in another year or so, if Mrs. Cole and Mr. Utting will continue their researches, I shall be able to answer in the affirmative.

Mr. DURBIN: I have a special interest in this project on the Social Accounts of Cambridgeshire because I had the good fortune while working at Cambridge to take part in some of the early discussions on it. As things turned out, I left before the project was under way and I made no serious contribution to it. However, I continue to watch its fortunes with keen interest.

My own concern in this work has been mainly on the sampling side, but before referring to that I should like to stress an important point which has not been emphasized so far, and that is the value of work of this kind in the collection of economic data for general economic analysis. Anybody who has worked in this field knows the difficulty of using data from administrative sources for econometric fitting and testing. In the last few years there has been a considerable development of elaborate statistical methods which have been very interesting and useful so far

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

as they go, but the results obtained from using them have been disappointing. The main reason, apart from defects in the theoretical models, has been the deficiencies of the data. The main hope for the future in such work lies in the use of direct methods of collection and I am therefore glad to welcome the contribution which the authors have made to the development of such methods. Their own work has been so far mainly in the field of family budget studies and it has been paralleled by similar work at Oxford and in the United States. However, it is worth noting that there are interesting possibilities ahead for applying methods of direct collection of data from business firms and other institutions.

On the sampling side I should like to comment on two aspects of sampling errors. First of all, the authors mentioned in a number of places in the paper the difficulty of obtaining estimates of the sampling errors of the aggregates in which they are interested. These aggregates are made up of a number of different items and for their method of calculating sampling errors, estimates of the co-variances of the different items are needed, the computation of which is an enormous burden. I should like to ask whether the authors have considered a different approach, namely, dividing a sample up into a number of random sub-samples and applying their methods of estimation separately to each of these sub-samples. The overall estimate of the aggregate would then be obtained as the mean of the estimates obtained from the different sub-samples. Hence the estimate of variance could be obtained by using the ordinary formula for calculating the variance of a mean. By dividing the sample up into ten parts, you would have an estimate of sampling error based on nine degrees of freedom; you could cut down or increase the number of sub-samples according to the importance you attach to having an accurate estimate of variance.

In discussions of errors in surveys a distinction is usually made between sampling errors and non-sampling errors. It is an important distinction and the authors referred to it a number of times in their paper. The sampling error is the part of the error due to the fact that only a fraction of the population is observed; the non-sampling error is the remainder of the total error. It is often said that the ordinary formulae for estimating variances provide estimates only of sampling error and do not allow at all for non-sampling error. This claim is incorrect since the estimate of variance depends on the differences between the observations on different individuals, and the observation on each individual contains a component of non-sampling error. Thus the estimate of variance will usually contain a component which depends on non-sampling errors.

By suitable design the sample can often be arranged so that a good deal of non-sampling error is allowed for in the estimate of variance. For example, a multi-stage design can be arranged so that each first-stage unit is dealt with by a different interviewer; hence a comparison between the means of different first-stage units will contain an amount depending on the differences between interviewers. Similarly, if it is felt that two questionnaires might be used in the survey, and one is uncertain which to choose, it is possible to arrange that both shall be used, randomizing the individuals to whom the questionnaires are given. The estimate of sampling error will then contain a component depending on the difference between questionnaires and, similarly, for time of day, weeks during which the survey is carried out, and so on. By using techniques of experimental design and in particular the device of randomization, it is possible to allow for a substantial part of the non-sampling error. However, such methods do nothing to reduce or allow for non-sampling bias. The problem of bias is a much more serious one which cannot be dealt with by simple statistical techniques.

Mr. MASON: I have not had time to study the paper deeply or to see its full implications, but one great merit is obviously that it enunciates many of the questions which have to be asked about household budget techniques.

A few more practical questions should be raised. If the interviewer goes on Friday he gets the "memory" of Thursday, but on Monday he usually gets Saturday's. (Interviewing on Sunday is very difficult). If this is so, may not the two day memory interval tend to exaggerate expenditure on the highest spending day?

Are the data for the monthly or longer periods taken from all households every week? If on only a section each time is there a possibility of bias occurring, especially in the fourth quarter of the year?

A big difficulty in some fields is to distinguish business expenditure from personal expenditure, as for instance the petrol used for the firm's business and that for the week-end jaunt.

Were wage bonuses at the end of the year correctly given? It is the "nett" which matters to the respondent; might there be a possibility of understatement of wages and salaries on that account?

The data on income may perhaps be 6-months lagged from that on consumption, as collected and aggregated.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Another element is the fortnight's holiday expenditure, which may amount to 8 per cent. of the expenditure for the year. I should welcome more details on this point. Household budget data would be more valuable sometimes if such holiday expenditure were separated.

Another valuable tabulation would be the provision of total amounts spent by the housewife, the children, including gifts, the head of the household and the subsidiary wage-earners, of which there are many to-day, especially in the Midlands. This type of data, even if it were not absolutely correct, would be very useful to those marketing a wide range of personal products.

Again the figures for, say, household purchases of paint and distemper may only be procured economically in this way; and even if it is shown only as "between £3–£5 per household", that is a useful basis to start from.

The more this kind of survey gets us facts not obtainable in other ways, even with all faults, the more we can improve business efficiency, and the more finance may become available to strengthen the statistical basis.

Could the hire purchase credit outstanding account for any of the discrepancy? This point has probably been well met; but a few more details on the general background would be useful. The raising of problems and the statement of practical solutions or present deficiencies make this a most valuable and practical paper.

Mr. UTTING (in reply): I am very grateful for many of the remarks which have been made, which will most certainly give us valuable lines on some of the things we had in mind to do with this fairly large body of data. In spite of its faults there is a good deal to be got out of it. I was particularly grateful for Mr. Durbin's suggestion about estimation of sampling errors by using sub-samples. We have, in fact, got all our material divided not only into quarters but into 13 four-week periods and tabulated in that way. Although this involves the element of seasonal variation it may give us reasonable estimates of some of these errors if we regard these as separate sub-samples.

There seems to have been a slight difficulty, largely because we only gave a brief summary in this paper, about just what this business of using different periods for different items involves. All the information, as we said, was collected by interview and therefore it *all* relates to past periods. It was collected as far as possible from separate individuals, seeing them separately, although this was not always possible in an agricultural cottage of one living room when it was pouring with rain in the winter. The information for some groups of items relates to the seven days immediately preceding the interview. There was no one-day information; it was always seven days. Similarly for the information which covered longer periods; it was the four weeks *immediately preceding* the interview and the year *immediately preceding* the interview. For some of the material which we have described as quarterly, such as electricity and gas, the bills are rendered quarterly and our information relates to the last bill received. This may refer to a period quite a long time earlier if the last bill was almost three full months ago and related to three months before that. This brings up the point made by Professor Allen about the time lag, which does have some influence on the size of our figures. It is only fair to say, not exactly in defence but as a partial explanation of this, that some of the expenditure items are likely to be understated for this reason, as well as some of the income.

Looking at this gap by various classifications other than income is something which we have very much in mind and are going on to. One further point vaguely related to something else we are going to do, is the treatment of the self-employed. Our Social Accounts of Cambridgeshire survey covers businesses as well as households and we are approaching a sample of self-employed from the business as well as the personal point of view. We shall have information from that survey which it may be possible to compare with the household side in some characteristics.

In thanking you for your reception of our paper, may I say that if anyone wants copies of our forms we will gladly send them to him.

The authors subsequently replied in writing, as follows:

Professor Allen thought us dogmatic in suggesting that record keeping is also subject to end-period effect. It is true that there is very little recently published evidence for this (the Ministry of Labour 1938 inquiry is probably the best known pre-war example); but we understand that some post-war record keeping inquiries in this country have obtained significantly higher estimates of expenditure from the first week of records than from subsequent weeks, and that this appears to be due to the importation of items of expenditure actually incurred before the beginning of the record keeping period. More information about this phenomenon is badly needed, and we agree with Professor Allen that it will be most interesting to see the results of the 1953 Ministry of Labour inquiry on this point.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Both Mr. Saunders and Mr. Kemsley have touched on the possibility of collecting all three items, income, saving and expenditure from each household. Our decision to split our sample was taken not so much because of concern with response as with the quality of the information supplied. This tends to deteriorate with an over-long interview. But if the three components are to be balanced on the spot for each household all three must relate to the same period, and this raises other difficulties.

Mr. Kemsley is right in supposing that there is appreciable variation in our sample numbers from week to week, and that these variations have not been balanced out within our quarters. It is true that in some circumstances, particularly if seasonal variation in the level of expenditure was marked by very sharp peaks or troughs, this could lead to a divergence between our last week's and period estimates. But in the case of clothing the period estimate itself refers to quite a short time (four weeks) and it is difficult to see how the weekly variations could lead to big differences unless the seasonalities were very sharp indeed.

We welcome the stress laid by Mr. Durbin on the contribution of non-sampling errors to the estimates of sampling error. But, of course, many of these non-sampling effects are biases, which cannot be made to operate randomly, and these may also be reflected in the estimate of sampling error—though in an undesirable form in this instance. A good illustration of this is that, if our hypothesis is correct and it is the lower income groups which overstate their expenditure on certain items, such as major durables, more than do the top income groups, then the variance of expenditure on these items will tend to be reduced. Our plea for more resources to be devoted to the problems we had mentioned was concerned more with bias of this kind than with error in a stricter statistical sense.

As a result of the ballot taken during the meeting the candidates named below were elected. Fellows of the Society:

Michael Aloysius Aczel
Harry Ball
George Dar
Alexander Geoffrey Ellinger
Ralph Alexis Vernon Fleming
Joseph Mark Gani
Robert Charles Gilder
Vidyadhar Prabhakar Godambe
Clive William John Granger
Pramod Kumar Gupta
Leonard Graham Hall
Ernest Alfred John Heath
George Roland Kendall
Gerald James Knight
Walter Lutz
Richard Henry Pain
Valerie Mavis Palmer
Peter Hanlon Richardson
Catherine Joan De Smith
Colin MacDonald Stewart
William Terence Thorneycroft.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Cost of Road Accidents

By D. J. Reynolds, B.Sc.(Econ.)

Road Research Laboratory, Department of Scientific and Industrial Research

Introduction

The occurrence of road accidents inflicts a burden on the community which may be considered in two parts.

(i) The pain, fear, and suffering imposed by the occurrence, or the risk of occurrence, of road accidents. These are considered of great importance in a society that values human life and human welfare.

(ii) The more concrete and ascertainable burdens in the form of the net loss of output of goods and services due to death and injury and the expenditure of resources necessary to make good the effects of accidents, e.g. medical expenses, vehicle repairs and costs of administration.

For a variety of reasons it is beyond the competence of the economist to assign objective values to the losses suffered under (i) and this paper is therefore confined to the estimation of the burdens listed under (ii). The estimates contained in this paper give, therefore, only the minimum values that can be placed on the consequence of road accidents and as such are not a complete guide to the policy to be adopted and the expenditure to be incurred on accident prevention.

In the following sections estimates will be made of the measurable cost to the community of road accidents occurring in 1952, together with estimates of the average costs of their various consequences, e.g. death, serious injury, slight injury, etc. By way of example these estimates are then applied to two accident-reducing measures to show how they may be used as a guide to the policy to be adopted and the action to be taken to prevent accidents.

Effects of Road Accidents

The effects of road accidents which are within the competence of the economist to evaluate are as follows:

(1) Damage to property, which consists mainly of damage to vehicles.
(2) Medical costs which are incurred for the treatment of casualties.
(3) Administrative costs which can be ascribed to the occurrence of accidents, e.g. the part of the administrative costs of motor insurance which can be ascribed to the occurrence of accidents.
(4) The net reduction in output of goods and services due to loss of output from people killed and injured, allowance being made in the case of persons killed for the goods and services they would otherwise have consumed.

In addition there are other effects which, although theoretically capable of evaluation, cannot be measured because of lack of data. These appear to be small and their exclusion will probably have very little effect on the final calculations. They are as follows:

(5) Administrative costs due to the occurrence of accidents, incurred by the police, the government and lawyers. These cannot be separated from the cost incurred during the performance of their other duties.
(6) The economic effects of a smaller population. With the existence of road accidents the population of Great Britain will be slightly smaller and of a slightly different age/sex occupational structure.
(7) The effect of transfers of income within the community. With the occurrence of accidents, income is transferred from the rest of the community in compensation to those who suffer loss from the accidents, without any necessary change in the resources at the disposal of the community. It is impossible to estimate this change.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

COST OF ROAD ACCIDENTS

In this section an attempt will be made to estimate how much the community is worse off in terms of measurable costs by the existence of road accidents. A certain difficulty arises here in that, whilst national income methods and data must be heavily relied upon in an approach of this kind, certain costs that will fall on the community are not taken into account and are not included in the national income, e.g. unpaid services. For example, if a housewife is killed or disabled so that her services are lost to her family, the national income is not affected even though the community has suffered a loss which (theoretically at least) is measurable. Also, when a man is killed the national income falls by an amount equal to the whole of his output, although it is clear that the community is worse off by the value of his net output only, i.e. his output less his consumption.

Since the loss of unpaid services and several other effects of accidents are an important part of the measurable loss borne by the community, they must be included in the total cost of road accidents, even though they are not considered to be part of the national income. For many reasons the resultant changes in the national income cannot be expected to reflect the extent to which the community is made worse off by the occurrence of road accidents, and the methods and conventions of national income accounting cannot be completely followed in this paper.

The cost to the community of road accidents is taken to be the costs incurred by items 1 to 4 (inclusive) above. These items are now discussed in more detail and estimates made of the costs or losses incurred by each item in 1952. Full details of the calculations on which the estimates are based are given below.

BASIS OF THE ESTIMATES OF THE COST OF ROAD ACCIDENTS IN 1952

All the figures given in this section are shown to be more significant figures than is justified by their accuracy. This is done solely to show the method of computation.

Cost of Damage to Property

An analysis of a sample of the records of an insurance company was made in September 1950 to discover the average cost of damage to property in 1949, per comprehensively insured car, goods vehicle and motorcycle. The results of this are shown in columns (b) and (c) of Table 1 The figures for public service vehicles were supplied by bus operators. The standard error of the total cost of damage caused by all vehicles except public service vehicles (£11·2 million) was £1·3 million. The remaining columns show how these figures are used to derive an estimate of the cost of damage to property in 1952, assuming that the total cost of damage to property will vary directly with the number of vehicles involved in personal-injury accidents in any year. The manner in which each column is derived is shown on the table. In column (h) the interim index of retail prices is used to correct the 1949 figures for changes in the value of money between 1949 and 1952.

Cost of Administration of Motor Insurance

The total amount spent in Great Britain on motor insurance has been estimated as £44 million in 1949 and £50 million in 1950 (Rudd, 1952). It can be estimated by similar methods to have been some £60 million in 1952. In 1949 and 1950, 39 per cent. and 38 per cent. respectively of the premiums paid to insurance companies were used for managerial and administrative purposes (Board of Trade, 1951). The expenditure on the administration of motor insurance is thus estimated to have been some £23 million in 1952.

Not all this expenditure, however, can be attributed to accidents, since even if few accidents occurred the need to insure against them would presumably remain, and as an arbitrary estimate it has been decided that one-half of the expenditure on administration may be attributed to road accidents. Costs of administration in 1952 are therefore estimated at some £12 million.

Cost of Medical Treatment

There is no direct information on the cost of the medical treatment of people injured in road accidents. The Social Survey estimated from its Survey of Sickness that 8 per cent. of days of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

TABLE 1

Cost of Damage to Property in Road Accidents: Great Britain: 1952

(a) *Type of Vehicles*	(b) *Average Cost per Vehicle of Damage to Vehicles* 1949 (£)	(c) *Average Cost per Vehicle of Damage to other Property* 1949 (£)	(d) *Number of Vehicles Registered* 1949 (*millions*)	(e) *Total Cost of Damage to Property* 1949 (£) *million* d (b + c)	(f) *Number of Vehicles Reported Involved in Personal-injury Accidents* 1949 (000)	(g) *Cost of Damage to Property per Vehicle Involved in Personal-injury Accidents* 1949 (£) (c/f)	(h) *Cost of Damage to Property per Vehicle Involved in Personal-injury Accidents* 1952 (£) (g×1·23)	(j) *Number of Vehicles Reported Involved in Personal-injury Accidents* 1952 (000)	(k) *Cost of Damage to Property* (£ *million*) (j, h)
Cars and taxis .	2·85	0·17	2·19	6·6	60	110	135	79	10·7
Goods vehicles .	3·22	0·46	0·84	3·1	42	73	90	46	4·1
Motor-cycles. .	2·02	0·28	0·65	1·49	29	51	63	46	2·9
All vehicles except Public Service Vehicles	—	—	—	11·2	—	—	—	—	17·7
Public Service Vehicles	4·2		0·7	0·30	27	12	15	27	0·4
All vehicles . .	—	—	—	11·5	—	—	—	—	18·1

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

incapacity and 8 per cent. of medical consultations in 1949 were the result of injuries of all kinds (Ministry of Health, 1951). The corresponding percentages in 1950 were 8 and 7 (Ministry of Health, 1952). It is known that about 12 per cent. of hospital beds are occupied by the chronically sick (Ministry of Health, 1952, Part II), so it is assumed that all injuries cost 8 per cent. of the remaining 88 per cent. of the cost of the hospitals, and 8 per cent. of the cost of the general medical and pharmaceutical services.

The only evidence on the proportion of those injuries resulting from road accidents is that given by Mackay (1951), namely the proportion of the victims of road accidents in all injured persons discharged from a sample of hospitals. (This proportion, 0·18, is not very different from the proportion of deaths in road accidents in all violent deaths, 0·22 (Registrar General, 1951).) On the basis of the above figures, it is assumed that road accidents accounted for 1·3 per cent. of the cost of the hospital services, and 1·4 per cent. of the cost of the general medical service and the administration of the health service in 1949 and 1950, as given in the annual reports of the Ministry of Health and the accounts of the National Health Service (1951, Parts I and II) to produce columns (a) and (c) of Table 2.

In Table 2 the average cost per person injured in 1949 and 1950 is calculated in terms of money of 1949 value, and the average for these years is used to estimate the cost of medical treatment in 1952. Changes in the value of money are adjusted using the interim index of retail prices. It is assumed that the injuries classified by the police as serious coincided with those receiving hospital treatment: this is a fair working approximation.

TABLE 2

Cost of the Medical Treatment of Persons Injured in Road Accidents: Great Britain

	1949		1950			1952	
			At 1950 Prices		*At 1949 Prices*		
	(a)	(b)	(c)	(d)	(e)		
	Total Cost (£000)	*Cost per Person Injured* (£)	*Total Cost* (£000)	*Cost per Person Injured* (£)	*Cost per Person Injured* (£) (d×0·9737)	*Cost per Person Injured* (£) $\left(\frac{b+e}{2}\times 1\cdot 23\right)$	*Total Cost* (£000)
Serious injuries:							
Hospital treatment	2,950	68·01	3,300	67·83	66·05	82·45	4,151
Administrative expenses	20	0·46	20	0·41	0·40	0·53	27
Total for serious injuries	2,970	68·47	3,320	68·24	66·45	82·98	4,178
Slight injuries:							
Treatment by general practitioners	760	5·92	800	5·42	5·28	6·89	1,054
Medical prescriptions	450	3·50	530	3·59	3·50	4·31	659
Administrative expenses	10	0·08	10	0·07	0·07	0·09	14
Total for slight injuries	1,220	9·50	1,340	9·08	8·85	11·29	1,727
Total all injuries	4,190	—	4,660	—	—	—	5,905

Reduction of Output

The first stage in evaluating losses of output was the estimation of average output per head (Tables 3 and 4). The value of the average annual output per head is taken to be the net domestic product at factor cost, divided by the number of persons in the working population. It is assumed that the outputs of men and of women are proportional to their earnings in the industries covered by the Ministry of Labour's enquiry (1953), and it is also assumed that the average output of housewives is equal to that of the average female worker.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 3

Estimates of the National Product of the United Kingdom in 1952

(£ million)

	(a) *Expenditure at Market Prices*	(b) *Net Taxes on Outlay (Including Local Rates)*	(c) *Expenditure at Factor Cost* (a — b)
Personal consumption	10,478	1,622	8,856
Public authorities' current expenditure on goods and services	2,915	74	1,841
Total consumption	13,393	1,696	11,697
Gross domestic capital formation	1,931	99	1,834
Exports less imports	195	71	124
Total taxes less subsidies	—	1,866	—
Gross national expenditure (= gross national product)	15,519	—	13,653
Net income from abroad	128	—	128
Gross domestic product	15,391	—	13,527
Depreciation allowances	704	—	704
Net domestic product	14,687	—	12,821

Source: Central Statistical Office (1953)

TABLE 4

Average Output and Average Consumption per Head: 1952

(a) Net domestic product at factor cost of United Kingdom (£ million)	12,821
(b) Of Great Britain (£ million)	12,472
(c) Working population of Great Britain (million persons)	
(i) men	15·86
(ii) women	7·28
(d) Weekly earnings (£)	
(i) men	8·8
(ii) women	4·7
(e) Proportionate shares of net domestic product	
$d_i,\ d_{ii}\left(\frac{b}{c_i d_i + c_{ii} d_{ii}}\right)$ = average output per head (£)	
(i) men	632
(ii) women	337
(f) Assumed output per head of housewives (£)	337
(g) Total consumption at factor cost (U.K.) (£ million)	11,697
(h) Population (U.K.) (million persons)	50·43
(i) Consumption per head $\left(\frac{g}{h}\right)$ (£)	232

Loss from deaths. The first stage in estimating the loss from deaths was to calculate the expected working life of people in each age group. This is done in Table 5 for men and in Table 6 for women. In both tables a part-time worker is counted as half a worker.

In Tables 7 and 8 the expected working life of each age group is multiplied by the average annual output per head, from which is deducted the annual average consumption per head to give the net loss of output due to the death of each member of each age group. This is multiplied by the number of deaths in each age group to give the total net loss from deaths in each age group. Since many of these losses lie well into the future they are discounted by an appropriate rate of interest over an appropriate period, to give the present value of the net loss of output due to deaths from accidents occurring in 1952. The results of these estimates are summarized in Table 9.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 5

Expected Working Life of Males, Great Britain, April–May, 1951

(a) *Age Last Birthday (years)*	(b) *Population* (000)	(c) *Total Occupied Population* (000)	(d) *Number Equivalent to Full-time Workers* (000)	(e) *Assumed Age Distribution of Workers* (000)	(f) *Average Number Surviving of* 1,000 *Live Births*	(g) *Average Number Working of* 1,000 *Live Births* $\left(£\frac{dc}{b}\right)$	(h) *Aggregate Future Working Years of Surviving Members of* 1,000 *Live Births*	(i) *Expected Working Life (years)* $\left(\frac{h}{f}\right)$
0–4	2,153	—	—	—	973	—	43,261	45
5–9	1,816	—	—	—	955	—	43,261	45
10–14	1,639	—	—	—	953	—	43,261	45
15–19	1,506	1,249	1,246	—	947	784	41,301	44
20–24	1,568	1,508	1,507	—	940	903	37,084	39
25–29	1,799	3,484	3,402	3,401	1,761	933	908	32,556	35
30–34	1,685				1,650	925	906	28,021	30
35–39	1,808	3,672	3,624	3,623	1,784	916	904	23,496	26
40–44	1,864				1,839	902	890	19,011	21
45–49	1,730	3,193	3,111	3,109	691	878	858	14,641	17
50–54	1,463				1,418	843	817	10,454	12
55–59	1,208	1,154	1,152	701	786	750	6,536	8
60–64	1,048	918	916		701	613	3,129	5
65–69	863	422	416	—	587	184	886	1·5
70 and over	1,309	273	263	—	—	—	88	—
Total	23,459	15,661	15,635	—	—	—	—	—

Source of columns (b) and (c): General Register Office and General Registry Office Scotland (1952).
Source of column (f): Registrar General (1952).

TABLE 7

Loss from the Death of Males in Accidents Occurring in 1952

(a) *Age Last Birthday (years)*	(b) *Average Expectation of Life (years)*	(c) *Average Expectation of Working Life*	(d) *Unrealized Output* (c×£632)	(e) *Unrealized Consumption* (b×£233)	(f) *Net Loss Per Male Killed*	(g) *Number Killed*	(h) *Total Net Loss* £000	(i) *Total Net Loss Discounted to Present* (1952)* £000
0–4	65	45	28,395	15,080	13,315	233	3,102	887
5–9	62	45	28,395	14,384	14,011	218	3,054	908
10–14	57	45	28,395	13,224	15,171	113	1,714	573
15–19	52	44	27,764	12,064	15,700	232	3,642	1,316
20–29	45	37	23,347	10,440	12,907	721	9,305	3,983
30–39	36	28	17,666	8,352	9,314	455	4,238	2,095
40–49	27	19	11,989	6,264	5,725	405	2,318	1,340
50–59	22	10	6,310	5,104	1,206	333	402	261
60–79	10	1·5	947	2,320	−1,373	701	−962	−791
80–	3	0	0	696	−696	149	−104	−96
						3,560	26,709	10,476

* The discounting factor is $\frac{1}{(1+R)^{\frac{n}{2}}}$ where R = the rate of interest (4 per cent. in this case) and n = the average expectation of life. For simplicity of calculation it is assumed that the whole of the loss due to the death of one person falls at the mid-point of the remainder of that person's life after accident.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

TABLE 6

Expected Working Life of Females: Great Britain, April–May 1951

(a) *Age Last Birthday (years)*	(b) *Population* (000)	(c) *Occupied Population* (000)	(d) *Of whom Part-time Employees* (000)	(e) *Other Part-time Employees* (000)	(f) *Total Number Equivalent to Full-time Workers* (000) $\left(c-\frac{d}{2}+\frac{e}{2}\right)$	(g) *Housewives (Full-time)* (000)	(h) *Housewives (Part-time)* (000)	(i) *Total Number Equivalent to Full-time Housewives* (000) $\left(g+\frac{h}{2}\frac{e}{2}\right)$	(j) *Number Surviving of* 1,000 *Live Births*	(k) *Number Working of* 1,000 *Live Births* (f, j)	(l) *Aggregate Future Working Years of those Surviving*	(m) *Average Expected Working Life (years)* $\left(\frac{l}{j}\right)$	(n) *Number of Housewives of* 1,000 *Live Births* (f, j)	(o) *Aggregate Future Years as a Housewife*	(p) *Average Expected Period as a Housewife (years)* $\left(\frac{o}{j}\right)$
0–4	2,054	—	—	—	—	—	—	—	976	—	18,453	19	—	24,390	25
5–9	1,735	—	—	—	—	—	—	—	964	—	18,453	19	—	24,390	25
10–14	1,585	—	—	—	—	—	—	—	962	—	18,453	19	—	24,390	25
15–19	1,570	1,227	5	—	1,225	41	1	42	958	747	16,586	17	25	24,328	25
20–24	1,678	1,110	35	39	1,112	501	30	497	953	632	13,138	14	282	23,560	25
25–29	1,837 } 3,582	1,326	166	131	1,309 { 758 }	1,162	159	2,176 { 1,026 }	946	390	10,583	11	528	21,535	23
30–34	1,745 }				551 }			1,150 }	938	296	8,868	9·5	618	18,670	20
35–39	1,883 } 3,787	1,303	244	198	1,280 { 602 }	2,380	232	2,397 { 1,245 }	928	297	7,386	8·0	618	15,580	17
40–44	1,904 }				678 }			1,152 }	918	327	5,826	6·3	555	12,648	14
45–49	1,812 } 3,466	1,183	194	166	1,169 { 704 }	1,150	176	2,155 { 1,093 }	902	350	4,133	4·6	544	9,900	11
50–54	1,654 }				465 }			1,062 }	878	247	2,641	3·0	564	7,130	8
55–59	1,481	410	60	48	398	986	50	907	843	2,227	1,456	1·7	562	4,315	5·1
60–64	1,329	191	30	20	186	977	23	979	791	1,111	611	0·7	582	1,455	1·8
65–69	1,148	103	15	—	95	—	—	—	710	559	186	0·3	—	—	—
70 and over	1,967	62	7	—	59	—	—	—	—	—	19	—	—	—	—
Total	25,382	6,915	756	602	6,833	—	—	9,153	—	—	—	—	—	—	—

Source of columns (b), (c), (d), (g) and (h): General Register Office and General Registry Office Scotland (1952). Source of column (j): Registrar General (1952).

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 8

Loss from the Death of Females in Accidents Occurring in 1952

(a) *Age Last Birthday (years)*	(b) *Average Expectation of Life (years)*	(c) *Average Expectation of Working Life (as Worker and Housewife)*	(d) *Unrealized Output* (c × £340)	(e) *Unrealized Consumption* (b × £233)	(f) *Net Loss Per Female Killed*	(g) *Number Killed*	(h) *Total Net Loss* £000	(i) *Total Net Loss Discounted to Present* (1952)* £000
0–4	69	44	14,960	16,008	−1,048	103	−107	−28
5–9	66	44	14,960	15,312	−352	98	−34	−7
10–14	61	44	14,960	14,152	808	31	25	8
15–19	56	42	14,280	12,992	1,288	62	105	35
20–29	49	36	2,240	11,368	872	101	88	34
30–39	40	27	9,180	9,280	−100	54	−5	−2
40–49	31	17·5	5,950	7,192	−1,242	76	−94	−50
50–59	22	9	3,060	5,104	−2,044	129	−263	−171
60–79	11	0	0	1,552	−2,552	372	−949	−750
80–	3	0	0	696	−696	113	−79	−73
						1,139	−1,313	−1,004

* The discounting factor is $\frac{1}{(1+R)^{\frac{n}{2}}}$ where R = the rate of interest (4 per cent. in this case) and n = the average expectation of life. For simplicity of calculation it is assumed that the whole of the loss due to the death of one person falls at the mid-point of the remainder of that person's life after accident.

TABLE 9

Net Loss of Output Due to Road Deaths Occurring in 1952

	£ *million*
Gross reduction in future output of goods and services	62
Future consumption of goods and services	36
Net loss of future output	26
Present values of net loss of future output	9·5

The loss of output from injuries.—To estimate the loss of output from injuries in road accidents it is necessary to know:

(i) How many people are prevented from working for any period as a result of injury in road accidents.
(ii) The period of time for which they are prevented from working.
(iii) How many people are still suffering from some disability when they return to work.
(iv) The duration of this further disability.
(v) The extent to which it reduces their output.

There is no direct information on any of these matters, and it was therefore necessary to make indirect estimates, using unpublished material supplied by the Ministry of National Insurance. Items (i) and (ii), the frequency and duration of spells of incapacity, were obtained for 1950 from sample analyses of the record of sickness benefit and industrial-injury benefit. These covered virtually all males of working age, and the first six months of incapacity from all injuries, whatever their cause. The proportion of all the loss of working time through injuries due to road accidents was assumed to be equal to the proportion of people injured in road accidents amongst those admitted to a sample of hospitals as a result of injuries (Mackay, 1951). As most women are not eligible for sickness benefit or injury benefit, it was assumed that they were disabled for a period which was in the same ratio to that of the men as women's admissions to the sample of hospitals were to men's.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

There are a number of minor objections to the results, some of which showed a need for small adjustments, but it seems likely that the final result is accurate to within ± 20 per cent. It is estimated that people of working age lost some 510,000 weeks' work in 1949 (males 370,000 weeks, females 140,000 weeks) in periods of disablement of under 6 months, and in the first six months of longer periods. Owing to the nature of the data, the estimate of loss from the remainder of the few long periods of disablement had to be included in the estimates of items (iii), (iv) and (v).

Information on items (iii), (iv) and (v), partial disablement for work, could be obtained only from the records of awards of disablement benefit which covered only people eligible for industrial-injury benefit. The data supplied by the Ministry of National Insurance showed the frequency distribution of degree of disability, as assessed for pension, for each type of injury listed in the *N* codes of the International Statistical Classification of Diseases, Injuries, and Causes of Death.

In using these data it was assumed:

First, that the number of disabled men eligible for disablement benefit was in the same ratio to the numbers ineligible as were the numbers of spells of sickness benefit resulting from injuries to the numbers of spells of industrial-injury benefit.

Secondly, that the proportion of disability, as assessed for the award of disablement benefit, was proportional to the reduction in the worker's output. This assumption had to be adopted in the absence of any better alternative. It seemed fairly satisfactory, except for very small assessments (under 10 per cent.) of permanent disability, where a large part of the assessment seemed to be for loss of general enjoyment rather than reduced ability to work. Accordingly half the calculated loss from very small permanent injuries was deducted.

Thirdly, it was assumed that the proportion of men disabled in road accidents amongst those suffering from each type of injury was equal to the proportion amongst men admitted to a sample of hospitals (Mackay, 1951) and suffering from the same type of injury.

Fourthly, it was assumed that the number disabled (a) of women of working age, (b) of girls of below working age, and (c) of males of below working age was in the same ratio to men of working age disabled as were persons injured in these groups to injured men of working age.

The duration of disability was also estimated from Ministry of National Insurance records, but this is subject to a particularly large margin of error, as, at the time of the analyses, many of the disabled had not been disabled sufficiently long for the ultimate duration of their disability to be known.

The police report each casualty as seriously injured or slightly injured. This classification has been taken as accurate.

Given the above assumptions, the estimation of the loss of output from injuries in road accidents in 1950 was simply a matter of arithmetic. Figures for 1952 were obtained by correcting the 1950 figures for changes in the numbers of persons injured.

Any estimate made in such a manner is bound to be subject to a large margin of error, which in this case may be as high as ± 50 per cent.

The final results obtained are summarized in Table 10.

TABLE 10

Loss of Output Due to Injury in Accidents occurring in 1952

	Serious Injuries	*Slight Injuries*
Number of persons injured	50,351	152,955
Loss of output in paid work (£ million)	20	4
Loss of housewives' services (£ million)	2	—
Total loss (£ million)	22	4
Total for all injuries (£ million)	26	
Possible error (£ million)	±13	

In Table 10 serious injuries (which are broadly those resulting in detention in hospital) and slight injuries are the classification of injuries used by the police.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TOTAL COST OF ROAD ACCIDENTS TO THE COMMUNITY IN 1952

Having estimated the individual items it is now possible to sum them, as in Table 11, to give the total cost to the community of road accidents occurring in 1952.

TABLE 11

Total Cost of Road Accidents occurring in 1952

Item	*Cost* (£ *million*)
Damage to property	18±4
Cost of medical treatment	6±2
Administrative cost of motor insurance	12±6
Net loss of output due to deaths	10±2
Loss of output due to injury	26±13
Total	72±15·0

The cost to the community of road accidents occurring in 1952 is therefore estimated at £72 million (± £15·0 million) equal to about 0·6 per cent. of the net domestic product at factor cost. If changes in price levels are taken into account this estimate is considerably lower than that made previously by Jones (1946).

Jones estimated the average annual cost of road accidents to the community between 1935 and 1938 at £60 million. Since these years the general level of prices, as given by the working-class cost of living index and the interim index of retail prices, has risen by some 83 per cent, which suggests that in terms of 1952 price levels Jones's estimate would be roughly £110 million, ie., some 50 per cent. higher than the above estimate.

This difference is at least partly explained by the following factors:

(i) The number of casualties from 1952 accidents was 208,012, of whom 4,706 were killed, whereas the annual average number of casualties in the years 1935–8 was 232,249 of whom 6,586 were killed (Jones, 1946). This difference in casualties between the two periods accounts for a difference of about £7 million in the two estimates.

(ii) Jones's estimate of the loss suffered by the community by death and injury was based on compensation awards made by courts. There is no reason to suppose that compensation awards, into which humanitarian considerations enter, should equal the economic loss actually suffered by the community and, in fact, they are likely to exceed it considerably. That humanitarian considerations are present is shown by the fact that for many injuries the compensation paid to females is the same as that paid to males (Jones, 1946), even though on the average a male worker's earnings and productivity are considerably higher than those of a female worker. The difference in approach, and in the evaluation of the economic loss suffered by the community, is further illustrated by the fact that in the estimate made by Jones no less than 90 per cent. of the cost of road accidents was attributed to human compensation for death and injury, compared with 50 per cent. in the estimate made above.

Since the estimates made by Jones are influenced by humanitarian considerations, which are specifically excluded from this paper, there is no reason to expect the two sets of estimates to agree at all closely.

THE COST OF VARIOUS CONSEQUENCES OF ROAD ACCIDENTS

To assess the economic gains to be obtained from particular schemes designed to reduce accidents it is necessary to know the cost of the various consequences of an accident, e.g. the average cost of a death, the average cost of a serious injury, etc. These are given in Table 12, the calculations on which the table is based being given in Appendix A.

A method of averaging over broad categories is used in Table 12. Although the cost of an accident falling within a particular category can vary widely with the persons and vehicles, etc. affected, a measure to prevent accidents would tend to reduce them indiscriminately and a method of averaging is appropriate; e.g. in the saving of a life, it is usually a matter of chance whether the life saved is of low or high economic value.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 12

Average Costs of Various Consequences of Accidents, 1952

Category	*Average cost* (£)
Death:	
All persons	2,000
Children (under 15)	3,000
Adults	1,800
Injury:	
Serious	520
Slight	40
Cost of damage to property per accident (injury and non-injury)	32
Administrative costs:	
Per injury accident	56
Per non-injury accident	6
Total cost per injury accident	332
Total cost per non-injury accident	38

An obvious exception to this is the death of children, who are especially valuable to the community and to whom differential safety measures can be applied, e.g. safety patrols, school propaganda. It is also possible to isolate the cost to the community of other particular types of accident or of accidents occurring to other road users to which special safety measures could be applied, but this kind of investigation belongs more to the special study than to a general analysis of accident costs and accident prevention.

Table 12 shows that the average cost of an accident or injury increases very sharply with its severity, e.g. an injury accident costs on the average about 9 times as much as a non-injury accident, whilst a serious-injury accident costs about 14 times as much as a slight injury accident, and a death costs more than 4 times as much as a serious-injury accident. The relative costs of each class of accident are probably not radically different, therefore, from the relative valuation which the community would be inclined to place on the avoidance of the different kinds of accident. It seems reasonably safe to assume, therefore, that from the viewpoint of the community the relative costs of different road accidents are roughly proportional to the suffering they cause, so that an economic calculation is a useful guide towards the relative priority which should be given to different safety measures. As an example of the use of these estimates as a guide to policy two examples of the economic return to be gained from road-safety measures are given in Appendix B. The examples are the introduction of new pedestrian crossings in 1951 and a proposal for an improved standard of rear lighting for vehicles. Both examples show a high rate of return on their capital costs, the return in the case of pedestrian crossings being very high.

CONCLUSION

Other examples of the economic assessment of road-safety measures could be given, including some on which, on strict economic grounds, expenditure could not be justified and others where the results are inconclusive. The examples given in Appendix B, however, serve to show that, at least in some cases, expenditure on road-safety measures not only yields a return in increasing human welfare but can lead to savings in real resources. An economic assessment which inevitably excludes humanitarian consideration can only give the very minimum return to be obtained from road-safety measures, so that it should not be relied on exclusively when deciding whether certain road-safety measures are justified. In particular, an economic assessment is not sufficient evidence to show that a particular road-safety measure should *not* be proceeded with. Nevertheless, the economic calculation does point to these measures which are likely to lead to savings in real resources and, in addition, increase human welfare and which, on these grounds, are, therefore, likely to be justified.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

SUMMARY

An estimate is made of the measurable cost to the community of road accidents occurring in 1952. The estimated cost, which includes cost of damage to property, medical expenses, the net loss of output due to death and injury, and administrative expenses, is put at about £72 million.

This figure is considerably less than that which would be obtained by adjusting estimates made by Professor Jones to present price levels, because his estimates were based on pre-war accidents which were more numerous than those occurring in 1952, and because some of his data were influenced by humanitarian considerations, which are specifically excluded from this assessment.

Estimates are also made of the average cost to the community of various results of accidents e.g. the average cost of a death in road accidents, in order that the economic justification for road-safety measures could be assessed. Two examples of the economic assessment of road-safety measures are given in an Appendix, the examples being the introduction of pedestrian crossings and improved rear lighting. The results show that both measures can be justified on purely economic grounds and *a fortiori* on humanitarian grounds.

ACKNOWLEDGMENT

The author wishes to acknowledge the work on costs of accidents carried out by Mr. Ernest Rudd of the Department of Scientific and Industrial Research, which has been the basis for much of the data contained in this paper. The paper is published by the permission of the Director of Road Research.

APPENDIX A

THE COSTS OF VARIOUS CONSEQUENCES OF ROAD ACCIDENTS

Cost of deaths.—The average cost of deaths of children (under 15) has been calculated by summing the net losses (discounted to the present) due to the death of children as shown in Tables 7 and 8 and dividing by the number of children killed in 1952. The average cost of deaths of adults is then the residual net loss from deaths divided by the number of adults killed in 1952.

Cost of injuries.—The costs of serious and slight injuries have been calculated by summing the costs of medical treatment and the loss of output attributable to each class of injury as given in Tables 2 and 10 and dividing by the number of injuries occurring in each class in 1952.

Cost of damage per accident.—In estimating the average cost of damage per accident a difficulty arises in that much of the damage to vehicles and other property takes place in accidents which involve damage only and which are not fully reported and recorded. In his study of the cost of road accidents, however, Professor Jones (1946) shows that the average cost of damage per non-injury accident in the years 1935 to 1938 was almost exactly the same as the average cost of damage per injury accident (£10·58 compared with £10·45). In the absence of evidence to the contrary, therefore, it seems safe to assume that the average cost of dmaage to property in damage-only accidents is the same as that in injury accidents.

Jones also shows that in a random sample of 2,555 accidents to Army vehicles occurring in 1942–3, where the tendency was to report even trivial accidents, 779 (30·5 per cent.) were injury accidents whilst 1,776 (69·5 per cent.) were non-injury accidents. This ratio of injury to non-injury accidents is considerably lower than that shown by police reports, e.g. in 1953 the ratio between injury and non-injury accidents in 25 English towns and the Metropolitan Police District was 37·63, but the tendency of the police to ignore minor non-injury accidents probably makes Jones's estimate the more accurate, even though it may be subject to some bias due to the class of vehicle and the conditions obtaining over the period covered by the sample.

If Jones's estimate is accepted, we may estimate that a total of about 572,000 accidents occurred in 1952, of which 171,757 were injury accidents and 400,000 non-injury accidents. Dividing the total cost of damage to property (£18 million) by the total number of accidents, the average cost of damage to property per accident (both injury and non-injury) is about £32.

Administrative costs.—As with costs of damage, administrative costs must be allocated between injury and non-injury accidents. The most reasonable assumption in this case seems to be that administrative costs will be distributed between the two categories according to their relative costs.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Ignoring administrative costs, of the £60 million total costs of accidents, non-injury accidents account for £12 million (20 per cent.) and injury accidents for £48 million (80 per cent.). On the above assumption, therefore, non-injury accidents account for 20 per cent. of administrative costs of £12 million giving a total of £2·4 million for 400,000 accidents, an average administrative cost of about £6. Injury accidents account for £9·6 million in administrative costs, an average of about £56 per accident.

Total cost for injury and non-injury accident.—For the sake of completeness estimates of the total costs of the average injury and non-injury accident have been given.

Revision of Estimates to take account of Changes in Price Levels

Owing to changes in price and wage levels the costs of the various consequences of accidents are likely to change over a period. It is difficult to choose a satisfactory index which will reflect the effect of price changes on the cost of accidents but it is suggested that, since the major components of accident costs are workers' incomes and labour costs, the estimates in Table 4 should be corrected for change in the Ministry of Labour Index of Weekly Wage Rates (Central Statistical Office) in order to bring them up to date in the future. The base date, corresponding to the date at which the estimates were calculated, may be taken as June 1952, at which the index (for all workers) stood at 129 (June 1947 = 100) and correction made for subsequent changes. As an example the index in June 1954 stood at 142 which suggests that the estimated costs in Table 4 should be increased by 10 per cent. in order to make them applicable to 1954 price and wage levels.

APPENDIX B

EXAMPLES OF THE ECONOMIC RETURN TO BE GAINED FROM ROAD-SAFETY MEASURES

The decision as to how much the community should spend on road-safety measures is considerably more complex than has been suggested above. In the first place the investigator is faced with a wide range of possible measures to reduce accidents, the efficiency of which he must test by a variety of experimental and statistical techniques. Because of the large number of factors which may contribute towards road accidents it is often difficult to isolate the effect of one particular measure, especially as in many cases the number of accidents both before and after the adoption of a particular measure may be so small that it is difficult to obtain statistically significant results. Also, because in many cases the number of accidents at the point of application of a road-safety measure (e.g. the installation of a particular pedestrian crossing) is small, it is necessary to consider the efficiency of the measure as a whole, so that it is not generally possible to carry out the complete marginal evaluation of the measure that a strict analysis would dictate. Conclusions on the effectiveness of safety measures and their economic justification therefore tend to be rather broad generalizations which will not necessarily apply to particular instances.

In the second place expenditure on safety measures may take a variety of forms and the benefits may be distributed in many different ways. Measures such as road improvements principally involve capital expenditure, whereas measures such as propaganda and safety patrols primarily involve current expenditure; the cost of the measure may be borne by the central and local authority (i.e. the tax-payer and rate-payer) or it may be borne by a particular class of road user, as when, for example, certain standards of equipment are imposed upon him. Similarly the benefits may be widely diffused and tend to benefit all road users or they may be limited to a particular class of road user.

In general, safety measures will tend to benefit the more vulnerable road user such as the pedestrian, cyclist and motor-cyclist. Since he will not necessarily bear the ultimate cost of the measure, an exact balancing of the costs of the measure against its resultant benefits for each class of road user cannot be expected, and to some extent the interests of the community and the interests of the individual road user may diverge.

The divergence between the benefits to the individual road user and the benefits to the community may prove an obstacle to the imposition of road-safety measures on particular classes of road users in a society in which individual freedom is valued. For example, as will be argued below, the imposition of higher standards of rear lighting appears likely to yield a clear benefit

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to the community, but to the individual road user, who may tend to ignore the probability of his being involved in an accident, it may seem a costly and unwarrantable interference with his liberty. The fact that a safety measure is likely to yield a clear benefit to the community, therefore, does not necessarily mean that the placing of various standards, restrictions and costs on the individual road user is justified.

The following examples have been chosen to show the method of assessing the economic return to be obtained from road-safety measures. For safety measures which involve capital investment, the decision as to whether the investment is justified on economic grounds may be made in the same way as any other investment decision. If the rate of return on the investment, measured by the resultant savings in accident costs, is greater than the marginal rate of return on investment in other directions (which may be assumed to be equal to the current rate of interest) and if the same result could not be obtained by any smaller expenditure, the case for the safety measure is clear. The investment will represent a net saving of resources, and on economic grounds and *a fortiori* on humanitarian grounds, will be justified. If the measure does not yield as much as the marginal rate of return on capital employed elsewhere, the final decision as to whether the measure is justified will depend on the valuation placed on human life and suffering by the community or the responsible authority.

Example 1.—*Pedestrian Crossings*

New pedestrian crossing regulations which came into force in October 1951 provided for a reduction in the number of uncontrolled crossings by about two-thirds and the striping of the remainder of the uncontrolled crossings.

A comparison of pedestrian casualties occurring in the first 10 months of 1952 and those occurring in the first 10 months of 1951 showed that pedestrian casualties fell by 8 per cent. between the two periods if allowance is made for the differences in the age structure of the population. The resultant annual savings in accident costs are as shown in Table 1B.

TABLE 1B

Reduction in Pedestrian Casualties 1951–52

Type of Casualty	*Reduction between 1951 and 1952*	*Savings in Accident Costs* £
Deaths:		
Pedestrian under 15 . . .	95	285,000
Pedestrian over 15 . . .	240	432,000
Serious injuries	1,563	813,000
Slight injuries	3,561	142,000
Total savings (£) . .		1,672,000

The estimation of the cost of carrying out the change in pedestrian crossing regulations presents some difficulties, chiefly because the cost of striping a crossing can vary widely with the methods used, e.g. whether a cheap process with a high maintenance cost or an initially expensive method with low maintenance cost is chosen. From returns from local authorities it is estimated that about 5,500 crossings were striped in 1952 and, choosing a method of striping which probably errs towards unusually high durability and high initial cost, the cost per crossing of a process with a life of at least 5 years would be £115. On this basis the cost of the whole scheme would be about £650,000 and depreciation and maintenance charges some £125,000 per annum.

If, therefore, the whole of the reduction in accident costs (£1,670,000) between 1951 and 1952 could be attributed to the new crossing regulations, this would represent a net return of some £1,550,000 per annum on an initial investment of £650,000, a yield of about 220 per cent. per annum. In fact, however, not all the reduction in pedestrian casualties between 1951 and 1952 can be attributed to the new crossing regulations since other factors such as improved street lighting, the increase in the number of television sets, and reduced economic activity, probably helped to reduce pedestrian casualties. Nevertheless, even if only one-fifth of the reduction in

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

pedestrian casualties is attributed to the permanent effect of the new crossings the net return on the investment would be some £185,000, equal to a yield of about 30 per cent. There is little doubt, therefore, that the measure can be justified on economic grounds. In this example the costs of removing disused crossings have been ignored, since these will tend to be off-set by decreased maintenance charges. Another factor which may be ignored is the effect of the new crossings on traffic flow. Available evidence indicates that, whilst delays to traffic at the individual crossings have increased, the reduction in the number of crossings has kept delays to traffic much the same as before.

Example 2.—*Rear Lights*

An estimate was made of the number of personal-injury accidents which may be attributed to inadequate rear lights on vehicles and cycles (Moore, 1952). It was found that the ratio of all accidents involving two vehicles (including cycles) occurring at night to those occurring by day in October to December 1949 was about 0·6, whereas the night/day ratio for certain types of accident which were closely dependent on adequate rear lighting at night (e.g. parked vehicle hit in rear) was significantly greater. The divergence between the expected number of these types of accident and the actual number occurring was attributed to inadequate rear lighting.

Allowing for the fact that police records show that many other accidents not usually associated with poor rear lighting are, in fact, caused by this factor, it was estimated that for a full year 2,800 accidents and 3,400 casualties would have been avoided if all vehicles and cycles exposed to risk at night had been fitted with adequate rear lights. If the distribution of these accidents between serious and slight injuries and deaths had been the same as the total number of night casualties occurring in 1950, their cost (at 1952 prices) would have been as shown in Table 2B.

TABLE 2B

Costs of Accidents Attributed to Inadequate Rear Lights in 1950

Item	*Number of Occurrences*	*Total Cost* (£)
Deaths:		
Children under 15	5	15,000
Adults	117	210,000
Injuries:		
Serious	962	500,000
Slight	2,316	93,000
Cost of damage per vehicle involved in accident	4,320	138,000
Administrative cost per personal-injury accident	2,800	157,000
	Total cost	1,113,000

Annual savings in accident costs of the order of £1,110,000 might have been expected therefore if adequate rear lights had been installed on all vehicles and cycles in 1950.

In 1950 a test of the rear lights of a small sample of vehicles showed that about 72 per cent. of private cars, 78 per cent. of commercial vehicles, 74 per cent. of public services vehicles, 71 per cent. of motor cycles and 98 per cent. of pedal cycles were below an adequate standard of rear lighting at that time, as defined by the American standard of intensity of 0·25 candela (Moore, 1952). It is, however, difficult to estimate the cost of bringing all vehicles and cycles up to this standard because very little is known of the cost of improving rear lighting of the individual vehicle. It is probable that motor vehicles could be brought up to the required standard without radical alteration at an average cost of perhaps £1 per vehicle, and tests at the Road Research Laboratory have shown that a cycle rear light could be brought up to standard by the use of the larger twin-cell battery lamp, at a cost of about 5*s.* per cycle. If these estimates are true the estimated cost in 1950 of bringing all rear lights up to an adequate standard would be as given in Table 3B.

In addition to this capital cost, additional maintenance costs would be incurred in maintaining the new standard of lighting. Additional maintenance costs for motor vehicles would probably

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 3B

Cost of Installing Rear Lighting of Adequate Intensity in 1950

Vehicle	*Number of Vehicles*	*Number Below Standard*	*Total Cost of Conversion to Standard* (£000)
Pedal cycles	6,000,000*	5,900,000	1,500
Private cars	2,258,000	1,626,000	1,626
Commercial vehicles	900,000	702,000	702
Public service vehicles	141,000	104,000	104
Motor-cycles	752,000	534,000	534
Other vehicles, ambulances, fire engines, etc.	67,000	50,000	50
		Total cost (£)	4,516

* Estimate of number of cycles in use in hours of darkness.

be negligible, but cyclists would incur additional expenditure on batteries. Tests on cycle lights suggest that the average cyclist with battery lights spends about 5*s*. 6*d*. a year on batteries (Rudd, 1952), and that about half of those using cycles after dark use battery lighting. If the cyclist using batteries increases his expenditure on them by about 50 per cent. the annual additional maintenance cost would be about £450,000, reducing the net return to be expected from the measure to about £650,000 per annum. On an investment of approximately £4½ million this represents a yield of about 14 per cent. on the initial investment. The margins of error of some of the estimates on which this calculation is based are considerable, particularly in the estimate of the cost of converting sub-standard vehicles and in the additional maintenance costs to be incurred, but the yield from the measure is sufficiently high to suggest that it would be worth making on economic grounds.

References

BOARD OF TRADE. (1951), *Assurance Companies acts* 1909–46. *Summary of Statements of assurance business deposited with the Board of Trade, London.* London: H.M.S.O.
CENTRAL STATISTICAL OFFICE. *Monthly Digest of Statistics.* London: H.M.S.O.
—— (1953), *National Income and Expenditures* 1946–1952. London: H.M.S.O.
GENERAL REGISTER OFFICE AND GENERAL REGISTRY OFFICE SCOTLAND (1952), *Census* 1951 *Great Britain One Per Cent. Sample Table, Part I.* London: H.M.S.O.
JONES, J. H. (1946), *Road Accidents. Report submitted to Minister of Transport.* London: H.M.S.O.
MACKAY, D. (1951), *Hospital Morbidity Statistics. Studies on Medical and Population Subjects No.* 4. General Register Office. London: H.M.S.O.
MINISTRY OF HEALTH (1951), *Report of* 1949–50 *Part I.* Cmd. 8342. London: H.M.S.O.
(1951), *Report for* 1949–50 *Part II.* Cmd. 8343. London: H.M.S.O.
—— (1952), *Report covering period for April* 1*st* 1950 *to December* 31*st* 1951, *Part II.* Cmd. 8582. London H.M.S.O.
MINISTRY OF LABOUR. (1953), *Monthly Digest of Statistics*, **87**, 120. London: H.M.S.O.
MOORE, R. L. (1952), *Rear Lights of Motor Vehicles and Pedal Cycles*, Department of Scientific and Industrial Research, Road Research Technical Paper No. 25. London: H.M.S.O.
REGISTRAR-GENERAL. (1951), *Statistical Review for England and Wales for* 1949. *Tables. Part I. Medical.* Now Annual Series, 29. London: H.M.S.O.
—— (1951), *Annual Abstract of Statistics*, 1938–50, **88**, 33. London: H.M.S.O.
RUDD, E. (1952), *J. R. Statist. Soc. A*, **115**, 179–198.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The New Zealand Agricultural Sample Survey

By W. B. Taylor

Applied Mathematics Laboratory, D.S.I.R., Wellington, N.Z.

AND

D. V. P. Clement

*Department of Census and Statistics, Wellington, N.Z.**

Introduction

For many years a census of numbers of livestock and acreage under crops has been taken annually for New Zealand as a whole and for its constituent areas. In 1949 a proposal was made to replace the annual agricultural and pastoral census by the collection of agricultural and pastoral statistics on a sample basis with provision for a complete enumeration every five years. It was realized that although such sampling methods would not provide the mass of statistical detail hitherto collected by the census, estimates of totals for the principal producing areas for each crop and of livestock totals could be obtained with known accuracy by this method.

In 1951 investigations proceeded at the Census and Statistics Department in consultation with the Applied Mathematics Laboratory, D.S.I.R., to determine the method to be used to obtain the most efficient estimate possible, with a given sample size of about 10 per cent. The most efficient sample design may be defined as that which, for a given sample size or cost of inquiry, gives the estimate of the quantities required with the smallest sampling errors. This paper gives the general outline of the design considerations of the sample inquiry which replaced the annual agricultural and pastoral census in 1953, together with a critical survey of the results of the sample collection.

Description of Material Covered

The area comprising New Zealand may, for the purposes of the survey, be divided into two groups according to the type of local administration; namely:

(i) Those areas known as "boroughs" administered by borough councils.

(ii) Other areas which fall under other types of local administration such as county councils, road boards, etc.

The land administered under borough control is mainly occupied by residences and commercial enterprises, and includes areas covered by cities and large towns and their environs.

The land lying outside borough boundaries, by far the larger in area, comprises the remaining land of the country from which are obtained all agricultural and pastoral products as well as raw materials, such as timber, coal, etc.

A county may or may not possess areas of both types, depending upon the existence within the county of sufficiently large towns involving the borough system, but all counties possess to some degree areas outside borough boundaries. Such areas vary widely between counties from 11,000 acres to over 1,500,000 acres and the "average" county contains some 200,000 acres, representing from between 400–700 holdings.

For the purposes of the census the unit of agricultural and pastoral production was taken as the "holding"—a holding consisting of any area of land one acre or more in extent situated outside borough boundaries. This was also used for sampling considerations.

The nature of the information to be collected by the sample survey followed closely that of the census. The major items included were:

* Since the publication of this paper the Department of Census and Statistics has been re-named the Department of Statistics.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(i) Land covered by the holding—with sub-divisions into Crown, Freehold and Leasehold ownership.

(ii) Areas and yields of field crops. Major questions concerned wheat, oats, barley, maize, peas, swedes, turnips, potatoes, flax, linseed, rape, and kale.

(iii) Areas in grass-seed, hay, and ensilage. Yields for all items in this category would be covered by the verification inquiry at a later date. Areas of grasses and clovers cut or to be cut in the season for hay or ensilage, as well as lucerne were also required.

(iv) Description of total area of holding. Areas in new pasture (laid down within the previous year and virgin land and land previously under cultivation). Areas in old pasture, tussock, fern, and scrub, standing timber, residence and out-buildings, and barren land useless for grazing and incapable of being made productive.

(v) Farm machinery at the time of inquiry: Number of rotary hoes and garden tractors, the number of agricultural tractors.

(vi) Livestock on the holding—subdivided into dairy stock, beef stock, pigs, horses, sheep and lambs, with age and sex subdivisions for dairy, beef stock, and pigs.

(vii) Areas of grassland top-dressed with artificial fertilizers and lime.

The county totals of agricultural and pastoral products vary widely because of the variation of the areas involved, the soil and pastoral conditions of each county and the intensity of farming practised. The larger counties possess a great proportion of unimproved or poor pasture land, such as mountain ranges and scrub country and, although such areas do not directly contribute in any great extent to the agricultural production of the country, the steady clearing of areas of unimproved land with their subsequent farming development, which is continuing, is an important factor in the general agricultural development of the country and consequently must be carefully watched.

DESIGN OF THE SURVEY

The raw data available for the design consisted of the returns of the 1950 and 1951 censuses as listed in the sub-enumerator books. The results of the 1950 census were also available on punched cards, each recorded holding covering six cards. Each holding of the 1950 census had been classified by the Department of Agriculture into one of 79 agricultural types upon consideration of the proportion of labour input involved in the principal types of farming (e.g. dairy, sheep, cattle). These data, although not forming an integral part of the sampling procedure, provided valuable information on the farming characteristics of each county and consequently facilitated the choice of the design employed. Maps used in the survey comprised geographical maps of New Zealand of 16 and 4 miles/inch, a soil map showing soil characteristics of the North Island from the Soil Bureau, D.S.I.R., and meteorological maps of rainfall, both annual and monthly averages, from the Meteorological Service of the Air Department. Pastoral vegetation maps from the Research Department and reports from the New Zealand Dairy Board were also used.

It was realized that while it would not be possible to provide estimates for all items for every county, estimates of county totals for those items for which the county was a principal producing area would be obtainable. In specific areas it would only be possible to estimate totals by "Land Districts", i.e. groups of adjacent counties, and for the North and South Islands. The information obtained with regard to the clearing and working of unimproved land could be supplemented by information available from the Department of Lands and Survey. Investigations were started to determine the possible accuracy of estimation of the major items of the inquiry obtainable from a sample comprising ten per cent. of the number of holdings. As estimates of county totals were required the county was taken as the domain of study and three counties, Hobson, Waipa, and Woodville, were chosen to obtain information as to the variability of the principal items of the survey. Some idea of the characteristics of these counties can be gained from Table 1 taken from the 1951 Agricultural census. Table 2 shows the distribution of holdings by area and by type of holding, a holding being classed as Mainly Dairy, for example, if 50 per cent. or more of the yearly labour input was devoted to dairy farming.

Both Hobson and Waipa represent fairly large counties, Hobson being predominantly dairying in character as is seen in Table 2. Waipa possesses greater sheep farming areas and a much smaller average size of holding. Woodville, on the other hand, is an altogether smaller county and was

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 1

1951–52 *Census Statistics for Hobson, Waipa and Woodville Counties*

	Hobson	*Waipa*	*Woodville*
Total number of holdings	1,166	1,809	361
Total area occupied (acres)	352,693	249,957	88,321
Average area of holding (ac.)	302	138	245
Area in grain and seed crops (ac.)	331	366	207
Areas in green fodder, root crops (ac.)	493	4,306	1,658
Area of grasses cut for seed, hay (ac.)	5,014	38,058	2,243
Number of cows in milk at January 31st	41,196	84,747	11,048
Number of sheep shorn (in season)	120,336	284,451	142,110
Number of pigs	14,033	21,703	3,043

TABLE 2

2*a*. *Distribution of Holdings by Area Groups.* 2*b*. *Distribution of holdings by Type Expressed as a Percentage of Total Number of Holdings, for Hobson, Waipa and Woodville* (1950 *Census*)

2*a*. *Area*

Acres	*Hobson*	*Waipa*	*Woodville*
1–10	52	186	34
10–40	95	181	36
40–100	426	934	142
100–250	317	369	71
250–500	163	109	45
500–750	56	16	16
750–1,000	42	12	14
1,000–2,000	7	2	3
2,000	8	—	—
Total	1,166	1,809	361

2*b*. *Type*

	Hobson	*Waipa*	*Woodville*
I. Mainly dairy	64·77	76·88	64·71
II. „ sheep	6·91	10·48	27·73
III. „ beef cattle	3·85	0·44	0·84
IV. „ arable cropping	—	—	—
V. „ mixed (including)	0·44	0·27	0·28
Poultry	0·09	0·60	—
Horses	—	0·98	—
Pigs	—	0·22	—
Total	0·70	2·93	1·12
VI. Other holdings	23·77	9·29	5·60

chosen in order to examine the coverage obtainable on such a county with a 10 per cent. sample.

Initially a one-stage sampling scheme was envisaged, a random sample being taken from each county. Later investigations showed that to obtain reasonable accuracy as well as adequate representation of all the holdings forming the county populations, stratification was also necessary.

Four sample designs each involving 10 per cent. of the number of holdings within each county were investigated for the three trial counties.

(i) A random sample of 10 per cent. of the holdings of each county.
(ii) A ten per cent. random sample stratified before selection by area of holding.
(iii) A random sample stratified before selection by agricultural "type" of holding.
(iv) A random sample involving stratification both by area and by type of holding.

The following well-known forms of estimation were considered:

(i) For the random sample

(*a*) The sample total in each instance is multiplied by the reciprocal of the sampling fraction (s.f.).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(*b*) Supplementary information in the form of the previous recorded census return for each sample unit is obtained and the ratio estimate (Yates, 1949), formed.

(ii) For the stratified samples

(*a*) For each item the sum of the weighted sample strata totals (w.s.t.), is formed, the weight being the reciprocal of the sampling fractions.

(*b*) Supplementary information in the form of the previous recorded census returns for each sample unit is obtained and the constant ratio estimate (c.r.e.) formed.

For the stratified sample both constant sampling fractions (c.s.f.) and variable sampling fractions (v.s.f.) were considered, the sampling fractions in the variable case being taken as proportional to the within-strata standard deviations for each particular item. Formulae for these estimates and expressions for their errors have been given by many authors (Yates, 1949 ; Deming, 1950 ; and Hansen, Hurwitz and Madow, 1953).

Stratification by type of holding proved unsatisfactory from the outset of the investigation. Although each holding of the 1950 census had been classified as belonging to one of 79 agricultural

TABLE 3

Cows in Milk at 31*st January* (1950 *and* 1951) *Hobson County*

		Area Group					
		1–249 *Acres*	250–749 *Acres*	750–999 *Acres*	1,000–1,999 *Acres*	2,000 *Acres*	*All Areas*
Type I	$S(x)$	23,775	12,140	868	1,394	77	38,254
	$S(y)$	24,149	12,305	905	1,513	81	38,953
	$V(x)$	518·7	922·0	522·7	5,484·7	1,502·0	841·1
	$V(y)$	468·7	1,001·1	481·6	11,378·3	1,301·1	1,012·0
	ρ	0·989	0·884	0·906	0·882	1·000	0·916
	n	529	183	13	12	2	739
Type II	$S(x)$	61	40	14	41	149	305
	$S(y)$	51	125	16	45	135	372
	$V(x)$	548·3	4·3	2·3	3·0	1,600·3	258·2
	$V(y)$	379·8	120·7	2·8	3·8	1,234·4	240·2
	ρ	0·958	− 0·064	0·804	0·645	0·999	0·883
	n	5	32	9	22	11	79
Type III	$S(x)$	10	10	5	7	0	32
	$S(y)$	3	131	32	7	1	174
	$V(x)$	2·7	1·9	1·0	12·25	0·0	1·9
	$V(y)$	0·2	302·2	93·5	12·25	0·0	135·3
	ρ	0·080	0·150	− 0·374	1·000	0·000	− 0·868
	n	16	17	6	4	1	44
Type V	$S(x)$	12	33	38	0	82	165
	$S(y)$	11	47	73	0	80	211
	$V(x)$	2·0	164·3	450·1	0·0	0·0	332·9
	$V(y)$	2·3	193·8	2,244·3	0·0	0·0	446·4
	ρ	0·958	0·915	1·000	0·000	0·000	0·939
	n	15	5	2	1	1	24
Type VI	$S(x)$	45	0	0	—	0	45
	$S(y)$	288	40	0	—	7	336
	$V(x)$	0·5	0·0	0·0	—	0·0	0·4
	$V(y)$	49·4	30·6	0·0	—	27·0	46·1
	ρ	0·044	0·000	0·000	—	0·000	0·041
	n	217	32	2	0	2	253
All types	$S(x)$	23,903	12,223	925	1,442	308	38,801
	$S(y)$	24,502	12,649	1,026	1,565	304	40,046
	$V(x)$	787·4	1,545·0	124·9	4,450·3	1,477·6	1,150·4
	$V(y)$	829·3	1,593·6	129·6	6,666·2	1,224·8	1,268·5
	ρ	0·956	0·939	0·965	0·938	0·995	0·949
	n	782	269	32	39	17	1,139

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

types of holding, the criterion involved in the classification had been a consideration of the labour involved on the holding rather than the type of agricultural production. The necessity for such a large number of types involved in the classification revealed the diverse nature of many holdings and the inherent difficulties in re-classifying the holdings into five or six "type" strata. It also appeared that many medium sized holdings could change in type within a short period of a few years and that such changes would depend on the relative economic values of such products as wool, mutton, beef, and butter over the particular period.

Stratification both by area and type was found unsatisfactory owing to the small number that would occur in the cross classification frame that would be necessary for sampling. An example of such numbers for Hobson County is shown in Table 3. From the 1950 census cards non-agricultural holdings were rejected and tables were prepared for the major items of the three counties. For Hobson these involved items such as: Dairy cows in milk; sheep (including stragglers) on holding shorn during the season just ended; total beef stock on holding; total grassland top-dressed with lime or artificial fertilizer; grasses and clovers to be cut this season for hay; and horses (all ages) on holding since the previous 31st January.

An example of one such table, Cows in Milk, is shown as Table 3. The actual tables used for examination of the various sampling schemes involved finer area groupings than that shown, the grouping being in 25-acre intervals up to holdings of 150 acres, then in 50-acre groupings to 749 acres, thence to 750–999 acres and 2,000 acres and above. Every holding in the county was taken from the 1950 census cards and the number of cows in milk recorded as the x variable. The cards were then matched with the enumerator books of the 1951 census and the corresponding y variable, the cows in milk for this year, was recorded. The variances $V(x)$, $V(y)$ and the correlation ρ between the two successive yearly figures were obtained for the holdings in each group. The subdivision by type of holding was also included, the six principal types of Table 2 being taken.

Consideration of the distribution of the area of holdings which, for all counties, followed a pattern similar to the counties examined, led to the introduction of points of stratification by area of holding at 250, 750, 1,000 and 2,000 acres. The accuracy of the methods of estimation given above for a 10 per cent. sample of the holdings of each county was then examined.

TABLE 4

Accuracy of Various Forms of Estimation in Hobson County

Calculated Standard Errors Expressed as a Percentage of Census Totals for a 10 per cent. Sample

Method of Estimation	*Cows in Milk*	*Total Area Top-dressed*	*Total Sheep Shorn*	*Total Pigs*	*Total Beef Stock*
Random sample . . .	9·0	13·0	30·0	103·6	37·0
w.s.t. Estimate:					
Area stn. const. s.f. . .	8·8	11·7	22·1	—	—
„ „ variable s.f. . .	6·0	8·2	19·3	—	—
Type stn. const. s.f. . . .	6·7	9·0	10·8	—	—
„ „ variable s.f. . .	2·2	6·2	3·4	—	—
Ratio estimation:					
Random sample . . .	2·7	8·0	18·6	48·5	13·2
Constant ratio area stn. . .	2·3	5·6	4·9	8·7	7·0
1952 Pilot survey . . .	3·0	5·9	5·3	9·8	8·3
1953 Sample survey result .	2·0	4·3	5·4	4·6	3·4
1952 Census total . . .	41,196	81,514	120,336	14,033	24,362

Table 4 shows the calculated standard errors as a percentage of the respective totals for the various sampling methods and forms of estimation. Stratification by areas of holding results in relatively smaller gains in accuracy for cows and top-dressed acreage than for sheep because of the lower correlation of these items with size of holding. Large holdings do not top-dress in the same proportion as the smaller more intensively farmed unit, and dairying, the most common type of farming for the smaller unit, forms only an optional smaller part of a large holding devoted to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sheep or cattle. Stratification by type of holding results in still increased accuracy, but as the sample allocation for each item was the optimum it was unlikely that a single multi-purpose sample could reproduce such accuracy. Ratio estimates, dependent on the current/historic correlation, showed a marked increase in accuracy for almost every item. Dairy herds are maintained annually by culling and replacement. Rotational top-dressing programmes occur on many large holdings. The change from pig raising to dried milk production which has been taking place in the dairy industry reduces the current/historic correlation for pig items.

On this evidence and for the reasons discussed earlier it was decided to stratify by area of holding with variable sampling fractions. Table 5 shows the within-strata standard errors for Hobson county for some of the items.

TABLE 5

Variances of Items by Area Groups—Hobson County

Item	1–249	250–749	750–999	1,000–1,999	2,000+
Cows in milk	829	1,594	130	6,666	1,225
Beef stock	14	495	2,565	3,776	25,220
Sheep shorn	1,625	45,540	228,946	364,262	2,041,772
Area top-dressed . . .	1,855	6,752	12,385	40,161	154,593
Total pigs	121	332	294	833	10,944
Total cattle	807	6,884	8,188	18,272	114,324

As expected, the size of the standard errors increases with area of holding for most items although the relative sizes differ widely from item to item. The ratio of variances from stratum to stratum differs for different items. Various possibilities for the choice of sampling fractions for multi-purpose surveys have been proposed. Sampling fractions proportional to average holding size within each strata, it was felt, would give too little weight to the smaller holdings. Deming (1950) suggests sampling proportional to strata numbers, i.e. constant s.f., suggesting that approximate maximum efficiency is nearly always obtained. When the cost of the survey is an important aspect, minimization of a suitable cost function for given values of the item variances provides a method of determining the s.f. However, solutions of the resulting equations are often difficult to obtain unless extremely simple cost functions are considered (Yates, 1953). In view of the diverse information required the question of adequate sample coverage was important and it was decided to consider a weighted linear function of the variances of the principal items of each county and to use such a function to determine the sampling fractions. Consequently the following scheme was introduced:

(i) For each county using the agricultural type tables together with the 1952 census estimates of the relative proportions of the New Zealand totals of the various items produced by the county, three items were chosen to represent the three principal types of farming practised in the area.

(ii) All holdings of 2,000 acres or greater in extent were included in the sample.

(iii) The remainder of the sample consisted of 10 per cent. of the total number of the remaining holdings within the county.

(iv) The number of holdings to be sampled from each of the four strata, 1–249, 250–749, 750–999, and 1,000–1,999 acres was determined by the formula

$$n_i = \frac{N N_i}{10}\left\{\frac{\sqrt{\left(\frac{s_{i1}^2}{T_1^2}+\frac{s_{i2}^2}{T_2^2}+\frac{s_{i3}^2}{T_3^2}\right)}}{\sum_{i=1}^{4} N_i\sqrt{\left(\frac{s_{i1}^2}{T_1^2}+\frac{s_{i2}^2}{T_2^2}+\frac{s_{i3}^2}{T_3^2}\right)}}\right\}$$

where n_i = sample number taken from the ith stratum.
N = total number of holdings within the four strata.
N_i = total number of holdings within the ith stratum.
T_j = total of jth item in the county (census previous) ($j = 1, 2, 3$).
s_{ij}^2 = variance of the jth item between holdings in the ith stratum.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

This formula was obtained by minimizing the function

$$\varphi = \frac{V_1}{T_1^2} + \frac{V_2}{T_2^2} + \frac{V_3}{T_3^2}$$

subject to $n_1 + n_2 + n_3 + n_4 = N/10$, where V_1, V_2, and V_3 denote the squares of the standard errors of the w.s.t. estimate from the stratified sample for the three items.

Although such a sample would contain about 13 per cent. of the total number of holdings, it provided for more accurate results than the 10 per cent. sample initially envisaged and permitted greater coverage of the areas involved. Calculations showed that 60 per cent. of the total acreage as recorded by the 1952 census would be covered by the inquiry, a figure deemed adequate for the survey.

Overall correlations computed for the three counties for the major items between successive years and also between the 1947–49 figures were in most cases high, and in view of this the use of supplementary information in the form of the census return of the year previous to the first sample survey was advocated. Correlations of the principal items for Waipa are given in Table 6. As strata totals were unobtainable from the census the constant ratio form of estimation was suggested.

TABLE 6

Correlations between Current and Historic Figures for Waipa and Hobson

Item	*Waipa* 1947 × 1948	*Waipa* 1947 × 1949	*Hobson* 1950 × 1951
Root crops	·415	·209	—
Hay	·647	·533	·722
Ensilage	·658	·449	—
Sown grass	·969	·904	—
Area top-dressed	·723	·675	·806
Electric motors	·635	·562	·861
Cows in milk	·883	·877	·949
Beef stock	·846	·657	·833
Sheep shorn	·955	·747	·882

Two pilot samples were taken using the above procedure. For Hobson, sampling the 1951 and 1952 census returns, estimates and their standard errors were computed, some of which are at the foot of Table 4. Using the 1951 census as a frame for sampling Woodville county, provisional estimates were made of the 1952 census before its complete tabulation. Greater attention was paid to the subdivisions of the major items, large errors being expected because of the small size of the county. Encouraging results were obtained (Table 7) on a sample of 36 holdings from the 361 recorded holdings within the county.

TABLE 7

Results of Woodville Pilot Sample Survey on 1951–52 *Census Data. Showing Estimates with Standard Deviations and Census Totals*

Item	*Sample Estimate*	*Census Total*
Total dairy cows and heifers	10,920±342	11,048
Heifer calves under 1 year	2,145±187	2,055
Beef cows and heifers 2 years and over	3,635±512	2,896
Total beef stock	9,312±370	9,925
Total cattle	25,577±521	26,087
Total pigs	2 521±462	3,043
Breeding bulls 2 years and over	297±39	347
Total horses (all ages)	549±32	587
Sheep shorn during last season	140,264±5,612	142,110
Lambs tailed during last season	97,316±1,939	95,683
Grasses and clovers cut for hay	2,092±211	2,018
Total grassland top-dressed	31,679±3,478	35,383

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Examination of the coverage of the sample in both cases showed a failing of the sample to obtain a reasonable cross section of the agricultural activities carried out on holdings of less than 250 acres. Market gardens, pig farms, orchards, and other types of holdings were inadequately represented, although it was realized that such agricultural practices occur principally in the dairying counties and counties adjacent to larger towns.

To obtain greater coverage over the smaller holdings it was decided to introduce a point of stratification at 100 acres. Following the procedure outlined above the sample number for the 1–249-acre group was obtained, this number was then divided into two numbers for the 1–99-acre and 100–249-acre groups the numbers being proportioned to the square roots of the total areas occupied by the holdings within these strata for the Land District in which the county was situated. It was decided, however, that where a county contained more than 1,000 holdings the numbers were to be taken as the square roots of the areas involved in the two strata for the county itself.

In the sample survey for some counties a further stratification of the 1–249-size group into 4-size groups of 1–9 acres, 10–39, 40–99, and 100–249 acres was carried out, the sample numbers being alloted as above. As the only complete list of holdings available consisted of the annual census, the last census (1952) before the commencement of the sample survey was taken as the basis for the frame of the sample. The sample chosen was expected to remain in operation for a five-year period or to the next census.

Data Collection

The data for the annual census were collected by personal canvass, the police acting as enumerators. This method led to accurate reporting, as the local police were well acquainted with residents of their district and cognisant of the need for accurate information. Such methods could not, however, be used for the survey, as previously the enumeration had been taken by the police in off duty time, and on a cost per schedule basis. The collection of information from only a fraction of the community spread over the same area rendered the method uneconomic from the collector's point of view. Consequently a postal collection was employed with a reminder and follow-up system where necessary.

Date and Duration of the Inquiry

Data to be collected consisted of the items mentioned above as they applied to all holdings on January 31st each year. Each collection and analysis was to occupy from 4–5 months annually from this date.

Schedule and Schedule Design

Schedules used for the previous census had been standard in design over the past years. It was felt that any change at this stage would re-introduce interpretation difficulties and confusion. Another difficulty in this sample inquiry was that an occupier upon receiving a schedule would ascertain that his neighbour had not had one and in these circumstances an unfamiliar schedule could lead to non-response or falsifications. For these reasons the survey schedule followed closely the pattern of its predecessors with some simplifying statements concerning some of the questions asked. Replacement of the census by a scientific selection of holdings was explained and accurate recording requested.

Preparation of Frame and Selection of Sampling Units

Three items were chosen, assessing each county separately and using the 1950 census returns on punched cards. The cards were sorted into the five area groups and the within strata variances computed. Three of the following were used to determine the sampling fractions:

(*a*) Sheep shorn on holding during last season.
(*b*) Total cattle on holding.
(*c*) Total dairy stock on holding.
(*d*) Area in grain and root crops.
(*e*) Number of cows in milk.
(*f*) Area top-dressed.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A card index of holdings was compiled for each county from the checked census. The holdings were then stratified into area groups, arranged alphabetically by name of occupier, and the sample was selected. Sampling fractions were determined as above. For a s.f. of $1/k$, say, a random start in the file from 1 to k was taken and thereafter every k^{th} holding was selected. The sample contained 11,845 of the 90,288 holdings of the frame and schedules were forwarded early in January, 1953. The numbers in each area group together with the initial response percentages in the subsequent two months is shown in Table 8.

TABLE 8

Response to Schedules of Sample Survey 1953
Mailed Early in January 1953

Area Group	1–99	100–249	250–749	750–999	1,000–1,999	2,000+	*Total*
Sample number . .	1,734	2,745	2,688	527	1,160	2,991	11,845
Percentages returned by:							
20.ii.53 . . .	54·7	54·2	56·0	51·4	48·5	49·9	52·9
6.iii.53 . . .	72·5	73·9	73·1	70·5	70·5	71·9	73·1
25.iii.53 . . .	90·0	91·0	91·3	90·3	91·1	92·1	91·2

Four reminders were sent at fortnightly periods and this resulted in a total final response of over 98 per cent. Some difficulties were experienced with area changes in the sampled holdings, mainly because of land transactions completed since the last census date, and adjustments were made in the calculations to allow for the different types of changes encountered, e.g. subdivisions, aggregations, and change to non-agricultural status.

Estimation of Results

One of three methods of estimation was employed in the sample for each county.

(i) *Constant Ratio* (Yates 1949)

This may be written as—

$$Y_R = \frac{X}{\Sigma g_i S_i(x)} \{\Sigma g_i S_i(y)\} = \frac{1}{c} \{\Sigma g_i S_i(y)\}$$

where $S(x)$, $S(y)$ denote the sample totals, the i suffix denoting the ith strata. Here y denotes the current sample value, x the previous census value returned by the holding and X the census total of that item for the county. This form of estimation was employed for values of $c \geqslant 0{\cdot}8$. Thus unless the "weighted strata estimate" of the base year from the sample

$$\Sigma g_i S_i(x)$$

estimated to within 20 per cent. of the census total the ratio method was not used. This prevented anomalous values of $1/c$ owing to possible inadequate coverage of an item by the sample in the base year.

(ii) *Difference Method* (Hansen, Hurwitz and Madow, 1953)—

For values of c between 50 and 80 per cent. (poor coverage) the item was estimated as:

$$Y_D = X + \Sigma g_i\{S_i(y) - S_i(x)\}$$

i.e. in the form of a correction to the existing previous census total.

(iii) *Weighted Strata Totals* (Yates, 1949)—

For lower values of c and virtually zero current/historic correlation the estimate

$$Y_W = \Sigma g_i S_i(y)$$

was used.

Estimates (ii) and (iii) were seldom necessary except for minor items, such as orchards, market gardens, grape vines, etc., and were not expected to yield very accurate results.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Estimates were computed for all items including major totals such as total dairy stock, total beef stock, etc., and the difference between the estimated total and the sum of the sub-item sample estimate divided between the component sub-items in each case in proportion to their estimated totals.

In calculating the variances and co-variances necessary to obtain the variances of the ratio estimates shown in Figs. 1–5, area changes were ignored. Any returned schedule showing a significant area change was regarded as having zero figures for every item of the schedule. This resulted in an inflated estimate of the component variances by reducing the correlation between the current returned figures and the base sample. Detailed attention to each estimate would provide greater accuracy for each item estimate than the methods so far described. Such examination, however, was reserved for special instances where the estimated value of the item showed an unusual rise or fall from the previous year's figure.

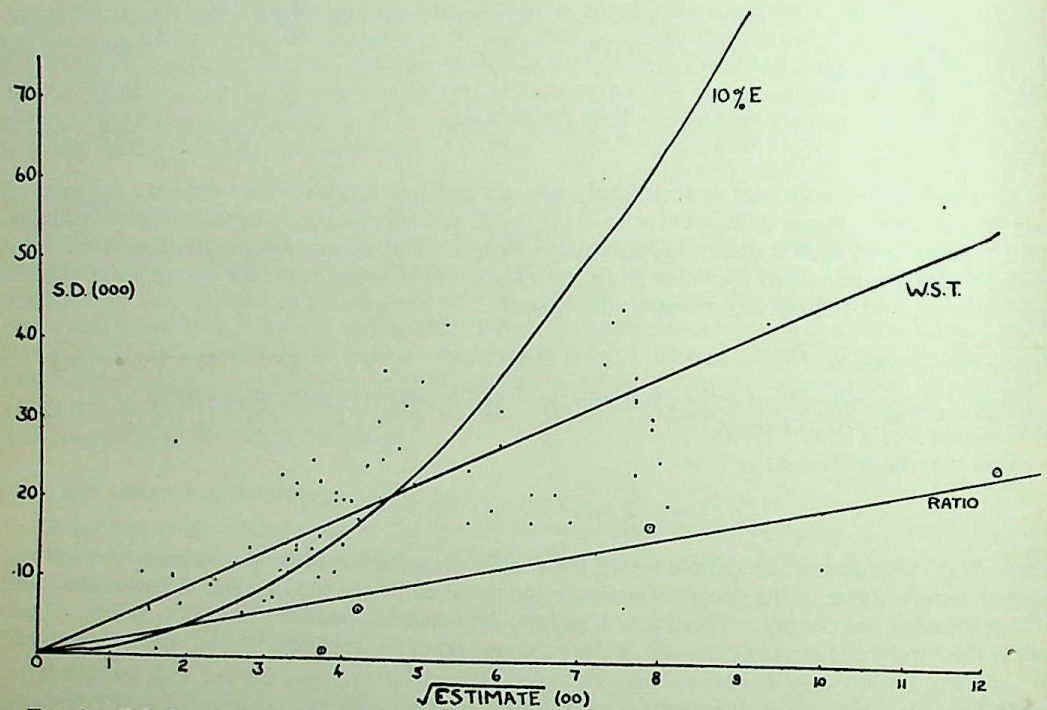


FIG. 1.—Calculated standard deviation of the estimate (S.D.) plotted against the square root of the estimate (County Totals) for the item "Number of Sheep Shorn in the Season".

From the within strata variances computed from the 1950 census of agriculture for those items used in determining the sampling fractions, it was possible to compute, for each county, the w.s.t. estimate for that year together with its variance based on the sampling fractions chosen and thus to obtain some idea of the possible accuracy of the current sample w.s.t. estimates forming the numerators of the ratios in the ratio form of estimation. Results showed that reasonable accuracy could be expected for most livestock estimates, the standard deviation in each case being approximately proportional to the square root of the estimate.

Fig. 1 shows the standard deviations plotted against the square root of the estimate for the 1950 census, the item being the total number of sheep shorn in the season; it illustrates the general relationship between estimate and standard deviation that appeared in most of the item calculations. Circled dots indicate results of the sample survey using the actual base year (1952) figures and the ratio method of estimation.

Every weighted strata total estimate computed, with a value exceeding 302,500, had a standard deviation well below 10 per cent. of the estimate, while few below this figure attained such accuracy. The computed ratio estimates, however (based on sample figures), had standard deviations well

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

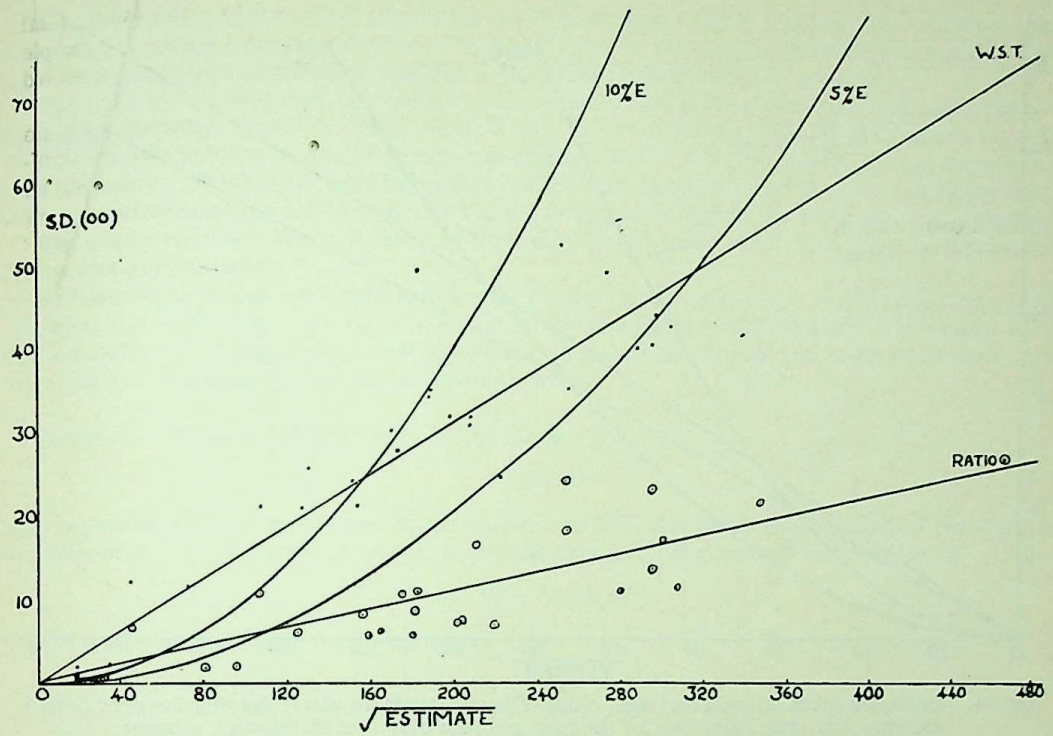


FIG. 2.—Standard deviation against square root of the estimated number of cows in milk.

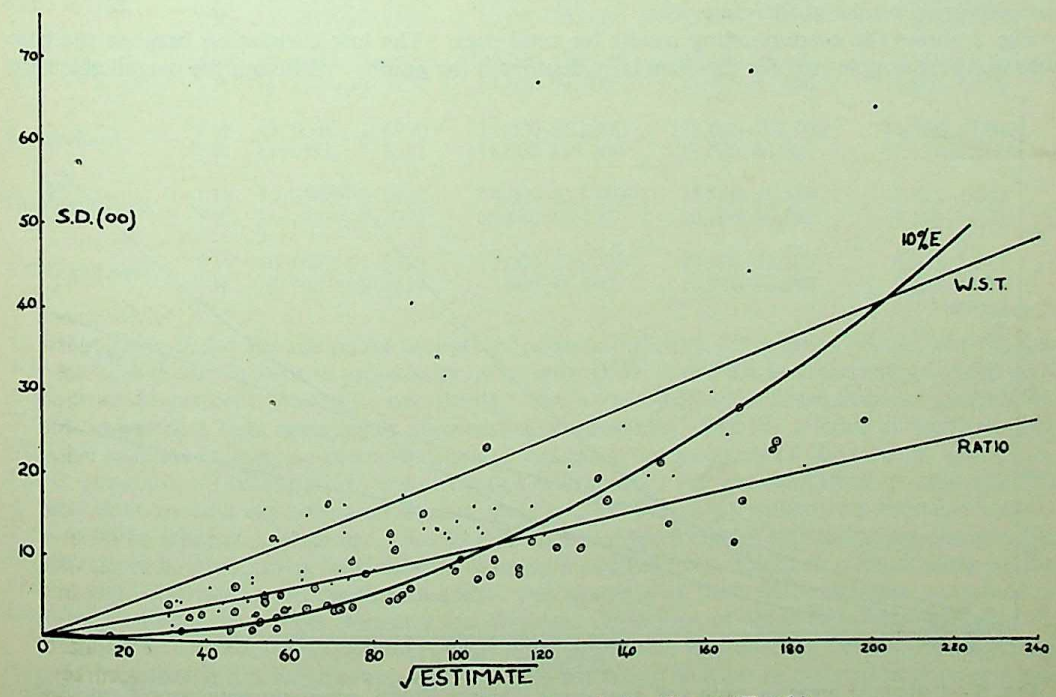


FIG. 3.—Estimated total number of pigs (County Totals).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

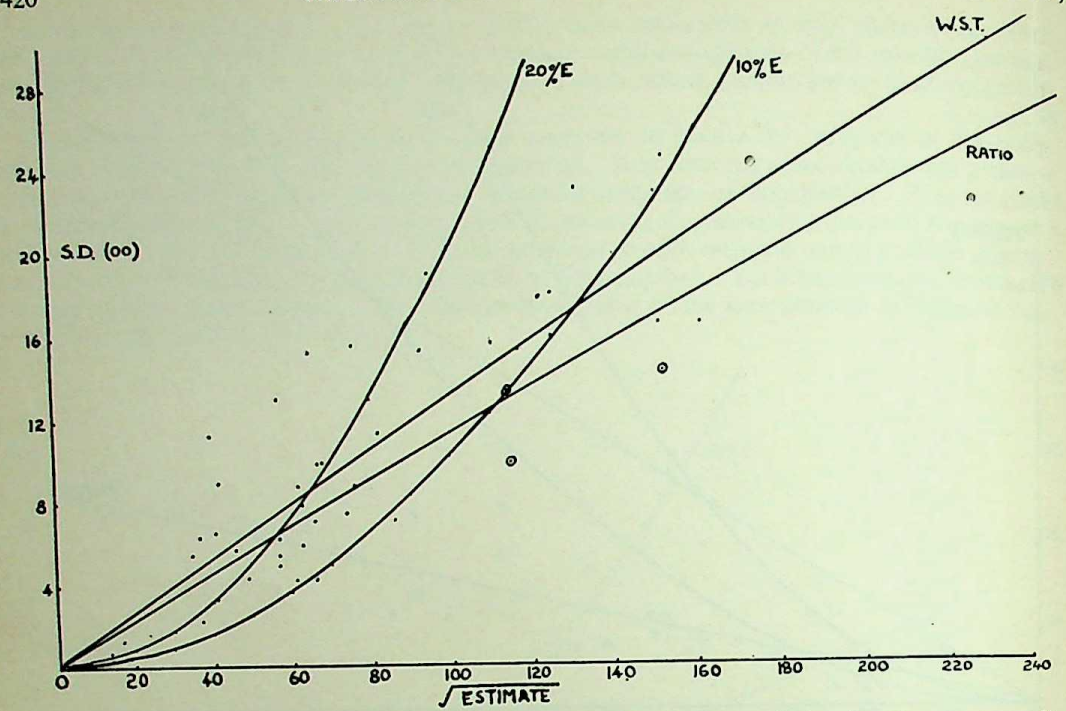


FIG. 4.—Estimated areas in grain and seed crops (County Totals, in acres) showing the poor overall accuracy and only slight gain of the ratio estimate over that of the w.s.t. estimate.

below 5 per cent. The results for cows in milk are shown in Fig. 2. In this case, the w.s.t. estimate was that obtained from the sample returns of the 1952 survey while the ratio estimates were the sample survey results of the same year.

Fig. 3 shows the corresponding results for total pigs. The low correlation between the base year and first sample year for this item is evident from the graph. Although the overall efficiency

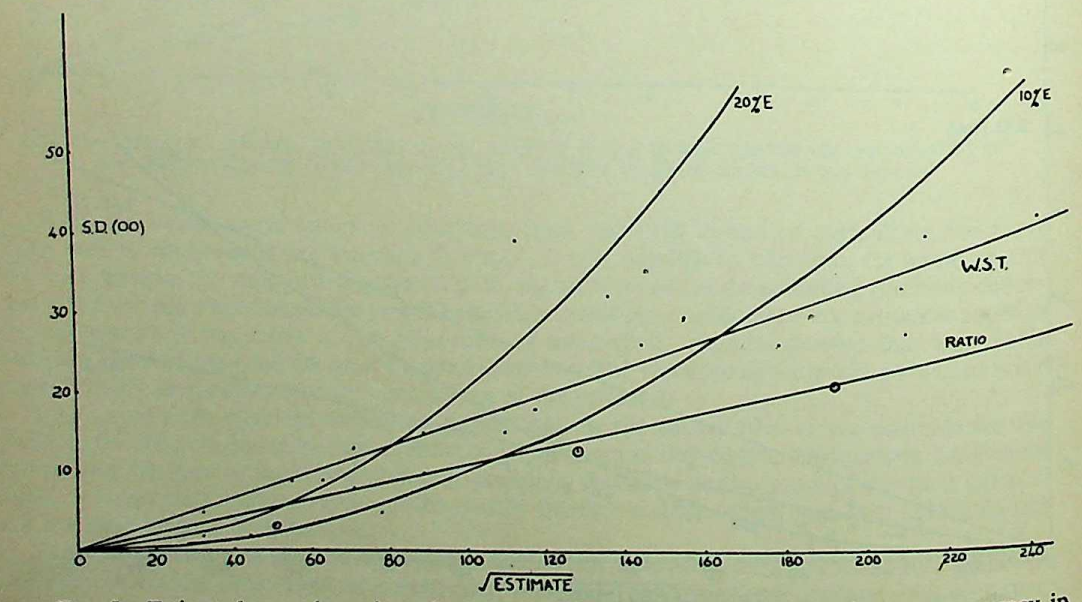


FIG. 5.—Estimated areas in grain and root crops (County Totals, acres) showing poor accuracy in both forms of estimation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the ratio method appears high (Table 10) the least squares lines in this case do not reflect true major trends, owing to the occurrence of high variances in some of the counties examined. The ratio form of estimation, however, resulted in an overall improvement in accuracy.

The inability of the survey to produce accurate yield figures for grain, root and seed crops is apparent from Figs. 4 and 5. One point in Fig. 4 is not shown, that of Ashburton County ($\sigma = 4{,}360$ $E = 160{,}720$ for the ratio estimate and $\sigma = 4{,}852$ $E = 162{,}463$ for the corresponding w.s.t. estimate). This county is an important one for such items, producing as it does 20 per cent. of the grain for threshing for New Zealand and a large percentage of the total pea and root crop.

The use of ratio estimation resulted in little gain in efficiency in most of these instances and in one or two cases a loss.

This low efficiency shown by the sample was detected at an early stage, as the collection of the schedules for Ashburton was one of the first to be completed. Subsequently a proposed verification inquiry was turned into a new sample survey collection for this and some other areas and more efficient estimates were obtained for these items.

Regression lines through the origin of the form

$$\sigma = b\sqrt{E}$$

have been fitted for the two forms of estimation and $S(b)$ the standard deviation of the slope b has been obtained. Values of b and of E corresponding to the points of intersection of

$$\sigma = [b \pm S(b)]\sqrt{E}$$

with $\sigma = 0{\cdot}20E$, $0{\cdot}10E$ and $0{\cdot}05E$ appear in Table 9.

TABLE 9

Showing Estimated Errors for Sizes of Various Estimates

Item	*Est.*	$b \pm S(b)$	5% E	10% E	20% E
Cows in milk	W	$15{\cdot}6200 \pm 0{\cdot}6584$	89,600–106,000	22,400–26,500	5,600–6,625
	R	$5{\cdot}5319 \pm 0{\cdot}3866$	10,600–14,000	2,650–3,500	662–875
Total pigs	W	$20{\cdot}1944 \pm 1{\cdot}4841$	140,000–188,000	35,000–47,000	8,750–11,750
	R	$11{\cdot}1344 \pm 0{\cdot}5409$	44,800–54,400	11,200–13,600	2,800–3,400
Sheep shorn	W	$43{\cdot}9046 \pm 1{\cdot}6990$	712,000–832,000	178,000–208,000	44,500–52,000
	R	$19{\cdot}4551 \pm 2{\cdot}5652$	114,000–178,000	28,500–44,500	7,125–11,125
Grain and seed	W	$14{\cdot}2745 \pm 2{\cdot}5542$	54,800–113,280	13,700–28,320	3,425–7,080
	R	$11{\cdot}1927 \pm 1{\cdot}1076$	40,800–60,400	10,200–15,100	2,550–3,775
Grain and root	W	$16{\cdot}6682 \pm 1{\cdot}0782$	97,200–126,000	24,300–31,500	6,075–7,875
	R	$10{\cdot}7487 \pm 0{\cdot}3154$	43,600–48,800	10,900–12,200	2,725–3,050

From Table 9, for the estimated number of cows in milk say, for a given county, with values between 22,400 and 26,500 the probability of its error (S.D. expressed as a percentage of the estimate) being 10 per cent. is roughly two-thirds. For a value below the lower limit the probability of attaining such a high accuracy is approximately one-sixth, while for a value above the upper limit the estimate is likely to be more accurate and fall to an accuracy of 10 per cent. with probability one-sixth. This applies to w.s.t. form of estimation. For constant ratio estimation 2,650 or more cows in milk are estimated as accurately as 22,400 by w.s.t. estimation, the same sample number being taken in each case. Overall efficiencies of the forms of estimation are comparable for each item by comparing their inverse regression coefficients. Variance computations of the various items have been proceeding steadily. An example of the counties already examined is provided by Rangatikei, some estimates and standard deviations of which are shown in Table 10.

As expected for most items showing a high correlation with the base year figures, the reduction in error by utilizing this information and the constant ratio (R) form of estimation is large when compared with the w.s.t. (W) estimate.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 10

Comparative Forms of Estimation from the Results of the Sample Survey in County Rangitikei

County Rangitikei (60)	1953 *Estimate*	*Standard Deviations*	*S.D.* (%)	*Estimation Method*
Freehold	542,674	7,746	1·4	R
	587,873	18,127	3·1	W
Total cattle	139,810	3,013	2·2	R
	146,408	5,296	3·6	W
Heifers (dairy) 1 year and under 2 years	3,404	508	14·9	R
	3,776	532	14·1	W
Area grassland *top dressed* (total)	156,827	10,482	6·7	R
	158,946	9,787	6·2	W
Lambs tailed on holding during season just ended	944,202	9,483	1·0	R
	970,011	33,675	3·5	W
Sheep Shorn during season and on holding at January 31st	1,489,541	24,958	1·7	R
	1,464,035	45,122	3·1	W
Grasses and clovers cut for *hay*	5,130	542	10·6	R
	5,520	1,266	22·9	W
Grain and root crops total	16,513	1,272	7·7	R
	14,500	1,285	8·90	W
Beef cows and heifers 2 years old and over	29,700	962	3·2	R
	30,715	2,024	6·6	W

The total figures for publication were obtained by summing the final individual county estimates. In 1954, in order to check the internal consistency of the survey estimates, land district and country (N.Z.) totals were also estimated by an alternative method.

This consisted in determining the overall ratio for the country as a whole by pooling the numerators and denominators of the individual county ratios and applying this factor to the N.Z. base year census totals. Again when $c < 0{\cdot}8$ a regression estimate was obtained of the form

$$\sum_{i,j} g_i S_i(y) + b\{ \sum_{j} X_j - \sum_{,j} g_i S_i(x)\}$$

where

$$b = \frac{\sum_{i,j} X_j g_i S_i(y)}{\sum_{i,j} X_j g_i S_i(x)}$$

The comparison between some of these estimates and those for publication is shown in Table 11. In view of the closeness of agreement between the results no changes were made in the conditions and methods of estimation.

CONCLUSIONS

The accuracy of most major items concerning livestock, areas top-dressed, in pasture, fallow and barren land, etc., proved adequate for the purposes of the survey and in all cases the use of supplementary information and ratio method of estimation was warranted in view of the gains in accuracy established. With crop-rotational and some smaller items, however, the use of ratio estimation resulted in little or no gain in accuracy when compared with the w.s.t. estimate and consideration of the sample returns has shown the need for a greater proportion of arable holdings within the sample.

Other points that became apparent were:

1. The 1–99-acre stratum, allocated by the method indicated in the design, resulted in too small a sampling fraction being chosen. Smaller, specialized holdings occurring within this group failed to attain adequate representation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 11

Sample Survey Estimates of New Zealand Farm Production Statistics (1954)

Item	*Published Figure*	*Check Figure % Difference*
Total area of holdings	42,359,588 acres	0·00
Areas in grain and root crops:		
Linseed for threshing	3,147 „	−3·70
Swedes	192,831 „	+0·22
Turnips (soft)	156,559 „	+0·29
Turnips (hard)	71,590 „	+2·35
Rape	176,051 „	+0·89
Kale	12,844 „	+7·08
Mangolds	2,010 „	+4·58
Potatoes	22,107 „	+0·33
Onions	1,343 „	−1·90
Pasture:		
New pasture (virgin land)	108,664 „	−0·03
Old pasture	16,937,484 „	+0·08
Areas top-dressed:		
Artificial fertilizers	5,337,238 „	−0·24
Lime	530,952 „	+3·13
Both	1,545,309 „	+1·12
Farm implements:		
Number rotary hoes and garden tractors	7,128	−15·06
Agricultural tractors	55,528	+0·21
Cattle:		
Dairy cows 2 years and above	1,995,018	−0·13
„ heifers 1 year and under 2	397,443	+0·06
„ „ under 2 years	481,457	−0·09
Beef cows, 1 year and above	28,212	+0·18
Total pigs	646,943	−0·64
Total horses	157,858	+0·17
Sheep shorn	35,472,952	+0·09
Lambs shorn	8,335,953	+0·08
Lambs tailed	33,558,048	+0·02

2. In view of the small number of holdings within the 750–999-acre group and the large variability within the 250–749-acre group, the point of stratification would appear to be better located at 500 acres.

3. Stratification of highly specialized holdings by type appears necessary in order to obtain information on such items as orchards, nurseries, vineyards.

4. More suitable forms of geographical stratification instead of the rather arbitrary administrative "county" should be examined. The large number of estimates required, many of which are based on relatively small sample numbers, increases the labour involved while tending to decrease the efficiency of the estimates.

5. The use of the "farm" as an economic unit to replace the "holding" would reduce the number of recorded area changes and the difficulty encountered in obtaining information from an owner on part only of his property.

6. Close and more detailed examination of land transactions through the offices of legal registration could simplify the technique of area changes in the sample holdings that were used.

7. Although the design principle of area stratification by size of holding with variable sampling fractions has been found satisfactory, further practical improvements are warranted. Investigations into area sampling and probability sampling would well repay the labour involved, especially in view of the low efficiency of the ratio form of estimates when estimating crop items. The low correlations between the current and previous season's returns are largely due to crop rotation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

It is hoped to carry out the sample survey together with a partial census linked to the Census of Population in 1956, the results of which will enable a further assessment of the accuracy of the survey to be obtained. First year results of the survey appear in the *Report on Farm Production Statistics of New Zealand* (1952–1953), published by the Department of Census and Statistics at Wellington, New Zealand.

ACKNOWLEDGMENT

We wish to thank the Government Statistician at the Department of Census and Statistics for permission to publish this paper.

References

DEMING, W. E. (1950) *Some Theory of Sampling*. New York: Wiley.
HANSEN, M., HURWITZ, W. & MADOW, W. (1953) *Sample Survey Methods and Theory*. New York: Wiley.
NEYMAN, J. (1934) "On the two different aspects of the representation method; the method of stratified sampling and the method of purposive selection", *J.R. Statist. Soc.*, **97**, 558–606.
YATES, F. (1949, 1953) *Sampling Methods for Censuses and Surveys*. 1st and 2nd ed. London: Griffin.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Some Theory of Index Numbers—II

Maintenance of Continuity by the Assumption of Proportionality of the New Index Number to the Old Index

By You Poh Seng,

University of Malaya

Introduction

A practice common among compilers and users of index numbers for joining a new series of index numbers to an old one is the following. An index constructed on base 0 = 100 undergoes one or more changes in, say, year T. In this latter year it is possible to obtain both the new index I'_T and the old index I_T. A continuous series is maintained from year 0 to year $(T-1)$ and onwards, incorporating the changes by a simple ratio process of the following form:

$$\frac{I'_T}{I_T} = \frac{I'_t}{I_t}, \quad t = 0, 1, 2, \ \ldots \ , T-1$$

so that

$$I'_t = \frac{I'_T}{I_T} \cdot I_t,$$

where I_t, like I_T, is the old index, and I'_t, like I'_T, is the new index which, if the proportionality postulated in the first equation above obtains, can be computed simply from the second equation.

There is no need to assume strict proportionality; so long as it holds approximately, I'_t can be computed approximately correctly. This fact will be kept in view in the discussion that follows.

For the sake of uniformity, the following notations will be used.

i = the original list of N commodities
$= 1, 2, \ \ldots \ , N.$

j = n added commodities, added at the end of the original list.
$= N+1, N+2, \ \ldots \ , N+n.$

k = p deleted commodities, assumed to affect the first p of the original commodities.
$= 1, 2, 3, \ \ldots \ , p.$

If, therefore, only addition is involved, the revised list of commodities will be

$$i' = 1, 2, \ \ldots \ , N, \ N+1, \ \ldots \ , \ N+n (= N+n \text{ items}).$$

If only deletion is involved

$$i' = p+1, p+2, \ \ldots \ , N (= N-p \text{ items})$$

If both addition and deletion of commodities are involved

$$i' = p+1, p+2, \ \ldots \ , N, N+1, N+2, \ \ldots \ , N+n (= N+n-p \text{ items})$$

As in my previous article in this *Journal* (Seng, 1956), we shall consider index numbers of prices with P denoting prices in general, and Q quantities.

Under each topic the discussion will deal with the various methods of index number construction in three major groups:

A.—The unweighted means of relatives:

(*a*) arithmetic mean; (*b*) geometric mean.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B.—The weighted means of relatives:

(*a*) arithmetic mean—weighted by:
  (i) base year weights.
  (ii) arbitrary year values.
  (iii) current year values.
(*b*) geometric mean—weighted similarly as above.

C.—Ratio-of-aggregatives, using as "basket of goods":

(i) base year quantities;
(ii) arbitrary year quantities;
(iii) current year quantities.

1. *Change of Base Year*

The method discussed in the previous paper in connection with the change in the base year is in fact the method of proportionality considered here. For the new index is based here on year T, and the old index on year 0, and we consider the proportion:

$$\frac{I_{tT}}{I_{t0}} = \frac{I_{TT}}{I_{T0}}$$

Since $I_{TT} = 100$, we consider in effect

$$I_{tT} = \frac{I_{t0}}{I_{T0}} \times 100$$

which was the method employed before. Since the conditions for the applicability of this method have already been considered in detail, we need not repeat them.

However it is interesting and useful to assess the error involved in applying the ratio conversion method for continuity for those index numbers for which the applicability of the method is conditional.

A. The unweighted mean of relatives

Only the arithmetic mean index number is to be studied here. We define

$$\frac{{}_tP_T}{{}_tP_0} = {}_ta,$$

and obtain

$$100\,\frac{I_{t0}}{I_{T0}} = \frac{100}{N\bar{a}}\,\Sigma\,\frac{{}_tP_T}{{}_tP_t}\ .\ {}_ta,$$

where $\bar{a}$ is the average overall commodities of the ratios of the prices in year T to those in year 0.

Thus the ratio process yields an index number of year t on base T, but it is an index number "weighted" by the proportions ${}_ta$. If these proportions are constant, then we obtain the required I_{tT}; otherwise the resulting index number will be biased in favour of those commodities whose prices in year T bear a high proportion to those in year 0.

The above equation can also be expressed as:

$$100\,\frac{I_{t0}}{I_{T0}} = \frac{100}{N}\,\Sigma\,\frac{{}_tP_t}{{}_tP_T} - \frac{100}{N}\,\Sigma\,\frac{{}_tP_t}{{}_tP_T}\left(1 - \frac{{}_ta}{\bar{a}}\right)$$

$$= I_{tT} - \frac{100}{N}\,\Sigma\,\frac{{}_tP_t}{{}_tP_T}\left(1 - \frac{{}_ta}{\bar{a}}\right).$$

And if we define ${}_tP_t/{}_tP_T = {}_t\rho$ and $\left(1 - \frac{{}_ta}{\bar{a}}\right) = {}_t\alpha$, then

$$100\,\frac{I_{t0}}{I_{T0}} = I_{tT} - 100\ \sigma_\rho\ \sigma_a\ r_{\rho\alpha}.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

That is, the error in the estimation involves the correlation between the price ratios P_t/P_T and the deviation from strict proportionality of ${}_iP_T$ to ${}_iP_0$. For we see that if there is strict proportionality of ${}_iP_T$ to ${}_iP_0$, ${}_ia/\bar{a} = 1$ for all i, so that $1 - {}_ia/\bar{a}$ designates departure from strict proportionality.

We also notice that whether the price-trend from year 0 to year T has been generally upwards or downwards, $r_{\rho\alpha}$ should be positive. This is equivalent to saying that, in general, the ratio process gives us an under-estimate of ${}^3I_{tT}$.

To assess approximately the magnitude of the error involved, we shall consider here only the probably more common case of rising price trend. This means that ${}_iP_t/{}_iP_T$ will in most cases be less than 1, say we take the range of ${}_i\rho$ for any one year t compared to T to be between 0·10 and 1·00. If we assume that the i commodities over the years can be taken as samples from a normal population, we can apply the findings of Tippett (1925) to estimate σ_ρ from the range. Suppose that the number of commodities N is 10, then the ratio range/σ is approximately 3, so that in the case being considered, $\sigma_\rho \doteqdot 0{\cdot}30$.

For departure from proportionality, we shall consider two cases:

(i) Where the departure is not extreme, say about 20 per cent. either way. Here ${}_i\alpha$ will range from about $-0{\cdot}2$ to 0·2, a difference of 0·4 from which we can estimate σ_α to be approximately 0·13. The under-estimate here is of the order of $4r_{\alpha\rho}$. The maximum under-estimate occurs when $r = 1$, and it decreases as r becomes smaller.

(ii) Where the departure is fairly large, α_i varying between $-0{\cdot}5$ and 0·5. Here σ_α is about 0·33, so that the under-estimate is $10r_{\alpha\rho}$, reaching a maximum magnitude of 10 as r approaches 1.

In the situation under discussion, the price trend is assumed to be one of increase, so that I_{tT} is less than 100. The errors above, therefore, when expressed in percentage forms, will be larger than in their absolute forms. If I_{tT} is say of the order of 50, the maximum under (i) will be 8 per cent., while that under (ii) will be 20 per cent.

We can therefore conclude that as far as departure of P_T/P_0 from proportionality is concerned, in the situation of rising price-trend and for the wide range of variations assumed for the price ratios P_t/P_T, probably the furthest departure from proportionality that can be allowed for P_T/P_0 is about 20 per cent. either way. If the number of commodities is larger, a wider difference can be tolerated. Also if the P_t/P_T's do not vary over a wide range, a similarly greater tolerance can be permitted.

As another example, if $N = 50$ and the P_t/P_T's vary from 0·5 to 1·0, then the maximum error for departures from proportionality of up to 100 per cent. either way is about 5, or 7 per cent., if we take for this particular case $I_{tT} \doteqdot 75$.

In respect of weighted index numbers, it can also be seen that the ratio process for changing the base year in most cases, in fact, yields the index number of year t to year T but with the weighting system affected to a certain extent. Thus:

(i) Weighted arithmetic mean, weighted by base year values. This is identical to the Lespeyre index number, where

$$100 \cdot \frac{I_{t0}}{I_{T0}} = 100 \cdot \frac{\Sigma_i P_t({}_iQ_T\ {}_ib)}{\Sigma_i P_T({}_iQ_T\ {}_ib)}, \qquad \left(\text{where } {}_ib = \frac{{}_iQ_0}{{}_iQ_T}\right) \cdot$$

$$= \frac{100}{\Sigma_i P_T\ {}_iQ_T\ {}_ib} \Sigma \frac{{}_iP_t}{{}_iP_T} \cdot {}_iP_T\ {}_iQ_T\ {}_ib.$$

(ii) Weighted arithmetic mean, weighted by current year values, where

$$100\,\frac{I_{t0}}{I_{T0}} = 100 \cdot \frac{\Sigma_i P_T\ {}_iQ_T}{\Sigma_i P_T\ {}_iQ_T\ {}_ic} \cdot \frac{1}{\Sigma_i P_t\ {}_iQ_t} \Sigma \frac{{}_iP_t}{{}_iP_T} \cdot {}_iP_t({}_iQ_t\ {}_ic),$$

$$\left(\text{where}^* \ {}_ic = \frac{{}_iP_T}{{}_iP_0}\right)$$

* For this index number, and that under (vi) the conditions for applicability of the ratio process are different from those given in the first article.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$= \frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_i P_t\, {}_iQ_t} \cdot \frac{\Sigma_i P_t\, {}_iQ_t\, {}_ic}{\Sigma_i P_T\, {}_iQ_T\, {}_ic} \cdot \frac{100}{\Sigma_i P_t\, {}_iQ_t\, {}_ic} \cdot \Sigma \frac{{}_iP_t}{{}_iP_T} \cdot {}_iP_t\, {}_iQ_t\, {}_ic.$$

(iii) Weighted arithmetic mean, weighted by arbitrary year values, where

$$100\,\frac{I_{t0}}{I_{T0}} = \frac{100}{\Sigma_i P_a({}_iQ_a\, {}_id)} \Sigma \frac{{}_iP_t}{{}_iP_T} \cdot {}_iP_a({}_iQ_a\, {}_id), \quad \left(\text{where } {}_id = \frac{{}_iP_T}{{}_iP_0}\right).$$

(iv) Weighted geometric mean, weighted by base year values, where

$$100\,\frac{I_{t0}}{I_{T0}} = 100^{\,\Sigma_i P_T\, {}_iQ_T\, {}_ie} \sqrt{\Pi \left(\frac{{}_iP_t}{{}_iP_T}\right)^{{}_iP_T\, {}_iQ_T\, {}_ie}}, \quad \left(\text{where } {}_ie = \frac{{}_iP_0\, {}_iQ_0}{{}_iP_T\, {}_iQ_T}\right).$$

(v) Weighted geometric mean, weighted by current year values, where

$$100\,\frac{I_{t0}}{I_{T0}} = 100\,\frac{\sqrt{\Pi \left(\frac{{}_iP_t}{{}_iP_T} \cdot {}_if\right)^{{}_iP_t\, {}_iQ_t}}}{\Sigma_i P_T\, {}_iQ_T \sqrt{\Pi({}_if)^{{}_iP_T\, {}_iQ_T}}}, \quad \left(\text{where } {}_if = \frac{{}_iP_T}{{}_iP_0}\right)$$

$$= \frac{\sqrt{\Pi({}_if)^{{}_iP_t\ {}_iQ_t}}}{\Sigma_i P_T\, {}_iQ_T \sqrt{\Pi({}_if)^{{}_iP_T\, {}_iQ_T}}} \cdot 100^{\,\Sigma_i P_t\, {}_iQ_t} \sqrt{\Pi \left(\frac{{}_iP_t}{{}_iP_T}\right)^{{}_iP_t\, {}_iQ_t}}$$

$$= \frac{\sqrt{\Pi({}_if)^{{}_iP_t\ {}_iQ_t}}}{\Sigma_i P_T\, {}_iQ_T \sqrt{\Pi({}_if)^{{}_iP_T\, {}_iQ_T}}} \cdot I_{tT}.$$

(vi) Paasche index number, where

$$100\,\frac{I_{t0}}{I_{T0}} = 100 \left(\frac{\Sigma_i P_t\ {}_iQ_t}{\Sigma_i P_T\, {}_iQ_T/{}_ih} \Big/ \frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_i P_T\, {}_iQ_T/{}_ih}\right), \quad \left(\text{where } {}_ih = \frac{{}_iP_T}{{}_iP_0}\right)$$

$$= 100 \left(\frac{\Sigma_i P_t\, {}_iQ_t}{\Sigma_i P_t\, {}_iQ_t/{}_ih} \Big/ \frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_i P_T\, {}_iQ_T/{}_ih}\right) \cdot \frac{\Sigma_i P_t\ {}_iQ_t/{}_ih}{\Sigma_i P_T\, {}_iQ_T/{}_ih}.$$

The discussion of the weighted arithmetic mean index numbers and of the ratio-of-aggregatives index numbers (which can be transformed to the weighted arithmetic mean forms) is therefore like that developed in respect of the unweighted arithmetic mean, and similar conclusions can be derived. The two geometric mean index numbers require some consideration.

We deal first with the base-weighted geometric mean form. Here

$$\log 100\,\frac{I_{t0}}{I_{T0}} = \frac{1}{\Sigma_i P_T\, {}_iQ_T\, {}_ie} \Sigma_i P_T\, {}_iQ_T\, {}_ie \cdot \log \frac{{}_iP_t}{{}_iP_T},$$

so that the problem is similar to that of the arithmetic mean, except that here we have to consider the logarithms of the price-ratios.

Finally, with regard to the geometric mean weighted by current year value weights, the ratio process in fact gives us the required index number I_{tT} with a factor which, if the value-weights of the commodities do not vary very violently from year to year, should normally be very near to unity. For this factor is the ratio of two "weighted" averages of the ${}_if$, the "weights" used being the values of year t and of year T.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2. *Addition of Commodities*

A.—The unweighted means of relatives:

(*a*) The arithmetic mean.

We consider the quantity

$$R = \frac{I'_{T0}}{I_{T0}} \Big/ \frac{I'_{t0}}{I_{t0}} = \frac{\dfrac{\dfrac{1}{N+n}\left(N \,.\, I_{T0} + 100 \sum_j \frac{{}_jP_T}{{}_jP_0}\right)}{I_{T0}}}{\dfrac{\dfrac{1}{N+n}\left(N \,.\, I_{t0} + 100 \sum_j \frac{{}_jP_t}{{}_jP_0}\right)}{I_{t0}}}$$

$$= \frac{N + \dfrac{100}{I_{T0}} \sum \frac{{}_jP_T}{{}_jP_0}}{N + \dfrac{100}{I_{t0}} \sum \frac{{}_jP_t}{{}_jP_0}}$$

$$= 1 + \frac{\dfrac{100}{I_{T0}} \sum \frac{{}_jP_T}{{}_jP_0} - \dfrac{100}{I_{t0}} \sum \frac{{}_jP_t}{{}_jP_0}}{N + \dfrac{100}{I_{t0}} \sum \frac{{}_jP_t}{{}_jP_0}}$$

$$= 1 + E \text{ (say)}.$$

If $E = 0$ for every year t, then the ratio process would yield precise estimates for

$$I'_{t0} \,(t = \text{year } 0, 1, 2, \quad \ldots \quad , T-1).$$

In general, of course, E would be 0 only in special circumstances, and we aim at studying the ranges of values for E to determine the degree of error involved in using the ratio process.

We introduce the concept of sub-indices here, and define

$$J_{T0} = \frac{100}{n} \sum \frac{{}_jP_T}{{}_jP_0}$$

and

$$J_{t0} = \frac{100}{n} \sum \frac{{}_jP_t}{{}_jP_0}$$

That is J is a sub-index relating only to the n added commodities

$$j(= N+1, N+2, \quad \ldots \quad , N+n).$$

We have then

$$E = \frac{\dfrac{J_{T0}}{I_{T0}} - \dfrac{J_{t0}}{I_{t0}}}{\dfrac{N}{n} + \dfrac{J_{t0}}{I_{t0}}}$$

E would therefore be zero for every year t only if J is proportionate to I for all years from 0 to T. This is a fairly unrealistic circumstance, and we must therefore consider the problem from a more realistic angle.

The magnitude of E depends on either the magnitude of $J_{T0}/I_{T0} - J_{t0}/I_{t0}$, or that of N/n, or that of both these quantities. If $J_{T0}/I_{T0} - J_{t0}/I_{t0}$ is small or if N/n is large, or if the former is fairly small and the latter fairly large, then E tends to be small.

Now we know the values of J_{T0}/I_{T0} and of N/n, but we do not know the value of J_{t0}/I_{t0}. Table 1 indicates the limits of J_{t0}/I_{t0} corresponding to various values of J_{T0}/I_{T0} and of N/n for an error (E) of $\pm$ 5 per cent. The figures in the top row refer to the minimum values of J_{t0}/I_{t0},

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and those in the bottom row to the maximum values, and it is not difficult to see that these limits yield errors of + 5 per cent. and − 5 per cent. respectively.

Since neither J_{T0}/I_{T0} nor J_{t0}/I_{t0} can be negative in value, it is possible that for certain combinations of J_{T0}/I_{T0} and N/n, J_{t0}/I_{t0} will have no lower positive values for $E = 5$ per cent.

As an example of the interpretation and use of the above table, we consider the following problem: An unweighted arithmetic mean index number, compiled on a coverage of 100 commodities from year 0 to year T (base 0) undergoes a revision in year T in the form of an addition of 5 commodities. In year T it is known that the average price ratio of the added commodities does not differ significantly from that of the original commodities. That is, $J_{T0}/I_{T0} = 1$; also $N/n = 20$. From the table the limits of J_{t0}/I_{t0} to yield errors not more than 5 per cent. range from the lowest possible value to around 2. In other words, if the prices of the added commodities are not known (as they generally are not) in any of the years $t = 0, 1, 2, \ldots, T-1$, but if we know approximately that the ratio which their general price level bears to the general price level

TABLE 1

Limits of J_{t0}/I_{t0} *corresponding to Various Values of* J_{T0}/I_{T0} *and of* N/n *for* E *to fall within the Range* ± 5 *per cent.* (*In the body of the table the figures on the top row denote the lower limits* [*corresponding to* + 5 *per cent. error*], *and those in the bottom row the upper limits* [*corresponding to* − 5 *per cent. error*]).

Ranges of J_{t0}/I_{t0} *for Corresponding* J_{T0}/I_{T0} *and for* N/n *as follows:*

J_{T0}/I_{T0}	10	20	30	40	50	60	70	80	90	100
0·1	a	a	a	a	a	a	a	a	a	a
	0·63	1·16	1·70	2·20	2·70	3·30	3·80	4·30	4·80	5·40
0·25	a	a	a	a	a	a	a	a	a	a
	0·8	1·3	1·8	2·4	2·9	3·4	4·0	4·5	5·0	5·5
0·5	a	a	a	a	a	a	a	a	a	a
	1·1	1·6	2·1	2·6	3·2	3·7	4·2	4·7	5·3	5·8
1·0	·5	a	a	a	a	a	a	a	a	a
	1·6	2·1	2·6	3·2	3·7	4·2	4·7	5·3	5·8	6·3
1·5	1·0	0·5	a	a	a	a	a	a	a	a
	2·1	2·6	3·2	3·7	4·2	4·7	5·3	5·8	6·3	6·8
2·0	1·4	1·0	0·5	a	a	a	a	a	a	a
	2·6	3·2	3·7	4·2	4·7	5·3	5·8	6·3	6·8	7·4
2·5	1·9	1·4	1·0	0·5	a	a	a	a	a	a
	3·2	3·7	4·2	4·7	5·3	5·8	6·3	6·8	7·4	7·9
3·0	2·4	1·9	1·4	1·0	0·5	a	a	a	a	a
	3·7	4·2	4·7	5·3	5·8	6·3	6·8	7·4	7·9	8·4
3·5	2·9	2·4	1·9	1·4	1·0	0·5	a	a	a	a
	4·2	4·7	5·3	5·8	6·3	6·8	7·4	7·9	8·4	9·0
4·0	3·3	2·9	2·4	1·9	1·4	1·0	0·5	a	a	a
	4·7	5·3	5·8	6·3	6·8	7·4	7·9	8·4	9·0	9·5
5·0	4·3	3·8	3·3	2·9	2·4	1·9	1·4	1·0	0·5	a
	5·8	6·3	6·8	7·4	7·9	8·4	9·0	9·5	10·0	10·5
7·5	6·7	6·2	5·7	5·2	4·8	4·3	3·8	3·3	2·9	2·4
	8·4	9·0	9·5	10·0	10·5	11·1	11·6	12·1	12·6	13·2
10·0	9·1	8·6	8·1	7·6	7·0	6·7	6·2	5·7	5·2	4·8
	11·1	11·6	12·1	12·6	13·2	13·7	14·2	14·7	15·3	15·8

(a) shows that there is no minimum positive value of J_t/I_t for the corresponding J_T/I_T and N/n.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the original list of commodities does not exceed 2 in any of these years, we can be certain that the ratio process of maintaining continuity will not lead us to error greater than 5 per cent. The nearer this ratio is to J_{T0}/I_{T0} (in this case 1), the smaller is the error.

If J_{t0}/I_{t0} were to exceed 2, the error increases fairly rapidly. Thus for $J_{t0}/I_{t0} = 3$, the error is around 9 per cent.; for $J_{t0}/I_{t0} = 5$, it is 16 per cent., and for $J_{t0}/I_{t0} = 10$, it is 30 per cent. On the other hand, if N/n were say of the order of 50, for the same value of J_{T0}/I_{T0} (= 1), J_{t0}/I_{t0} would have to be about 4 for an error of 5 per cent. to be reached, while a value of J_{t0}/I_{t0} of 10 would involve an error of 15 per cent.

If J_{T0}/I_{T0} were to be 2, then the limits of J_{t0}/I_{t0} are, for $N/n = 20$, between 1 and about 3, while for $N/n = 50$, there would be no lower limit, and the upper limit would be around 5.

Thus, although the prices of the added commodities from year 0 to year $T-1$ may not be known, an idea of the ratio which their general level bears to that of the original commodities enables us to estimate how accurate the ratio process of maintaining continuity is likely to be, and therefore whether the method is useful in any particular case.

(*b*) The geometric mean.

The quantity R is:

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = \frac{\frac{(I_{T0})^{\frac{N}{N+n}}}{I_{T0}} \cdot \left(\prod_j \frac{{}_jP_T}{{}_jP_0}\right)^{\frac{1}{N+n}}}{\frac{(I_{t0})^{\frac{N}{N+n}}}{I_{t0}} \cdot \left(\prod_j \frac{{}_jP_t}{{}_jP_0}\right)^{\frac{1}{N+n}}}$$

$$= \left[\frac{\frac{J_{T0}}{I_{T0}}}{\frac{J_{t0}}{I_{t0}}}\right]^{\frac{1}{\frac{N}{n}+1}}$$

where

$$J_{T0} = \sqrt[n]{\prod_j \frac{{}_jP_T}{{}_jP_0}}$$

and

$$J_{t0} = \sqrt[n]{\prod_j \frac{{}_jP_t}{{}_jP_0}}$$

and are sub-indices computed by the same method as the main indices.

As before if $R = 1$, then the ratio process is an exact method for maintaining continuity. Our aim is to estimate the deviation of R from 1, i.e., if we put $R = 1 \pm E$, to estimate the magnitude of E.

Table 2 shows the limits for 5 per cent. errors in R, of J_{t0}/I_{t0} for various values of J_{T0}/I_{T0} and N/n, these being the three quantities on which R depends. The set-up of the table, its interpretation and use are similar to those of Table 1.

A comparison of Table 2 with Table 1 shows that for most of the values of J_{T0}/I_{T0} and N/n, the ranges of J_{t0}/I_{t0} are much wider for the geometric mean than for the arithmetic mean. Thus for $J_{T0}/I_{T0} = 2$ and $N/n = 20$, the ranges of J_{t0}/I_{t0} for the arithmetic mean are 1·0 to 3·2, while those for the geometric mean are 0·72 and 5·88. Only for a small number of values of J_{T0}/I_{T0} (= 0·1 and 0·5), and then only for N/n of about 50, are the ranges narrower. That is to say, the geometric mean method of construction in general gives smaller errors than the arithmetic mean method of construction in the application of the ratio process of maintaining continuity.

Another important point in the table is that the value of N/n affects the error to a much greater extent than in the case of the arithmetic mean, particularly for largish values of N/n. Thus for $J_{T0}/I_{T0} = 1$ and $N/n = 30$, J_{t0}/I_{t0} has to be as small as 0·22 and as large as about 5 for the error to be 5 per cent.; for $N/n = 50$, the range is from ·08 to about 14. If now $J_{T0}/I_{T0} = 1$, $N/n = 50$ and $J_{t0}/I_{t0} = 3$ say, the error is only of the order of 2 per cent. while for the arithmetic mean it is about 4 per cent.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 2

Limits of J_{t_0}/I_{t_0} *corresponding to Various Values of* J_{T_0}/I_{T_0} *and of* N/n *for* E *to fall within the Range* ± 5 *per cent. (In the body of the table the figures in the top row denote the lower limits [corresponding to* + 5 *per cent. error] and those in the bottom row denote the upper limits [corresponding to* − 5 *per cent. error].)*

J_{T_0}/I_{T_0}	*Ranges of* J_{t_0}/I_{t_0} *for Corresponding* J_{T_0}/I_{T_0} *and for* N/n *as follows:* 10	20	30	50	100
0·1	0·06	0·04	0·02	0·008	0·001
	0·18	0·29	0·49	1·37	17·87
0·5	0·29	0·18	0·11	0·04	0·004
	0·88	1·47	2·46	6·86	89·4
1·0	0·58	0·36	0·22	0·08	0·007
	1·76	2·94	4·91	13·72	178·7
2·0	1·17	0·72	0·44	0·17	0·014
	3·52	5·88	9·82	27·44	357·6
				3·150	
3·0	1·76	1·08	0·66	0·25	0·02
	5·28	8·82	14·73	41·15	536·3
4·0	2·34	1·44	0·88	0·33	0·03
	7·04	11·76	19·65	54·88	715·2
5·0	2·92	1·80	1·10	0·41	0·036
	8·80	14·70	24·57	68·60	893·9
10·0	5·8	3·6	2·2	0·8	0·07
	17·6	29·4	49·1	137·2	1787·0

B.—The weighted means of relatives.

(*a*) The arithmetic mean.

(i) Weighted by base year values. In the discussion of this section, it is important to point out that we are concerned with the situation that in year T, when the index number is altered by the addition of the n new commodities, the prices of these commodities for years before year T are not available. It also follows that the values of the added commodities are not available for the base year 0. This being so, the change in the index-number compilation is likely to involve not only the addition of new commodities, but also a revision of the weighting system to a later year (possibly T), when the prices, and therefore the values, of the added commodities are available. This will involve also a shift of the base year.

It is therefore reasonable to conclude that in altering the index-number of this form by the addition of new commodities, two situations can arise, either (*a*) the prices and values of the added commodities are available for the base year 0, in which case the alteration would involve only the one aspect of addition of commodities, and the maintenance of continuity in the situation of availability of price-date has already been studied in the previous article; or (*b*) the prices and values of the added commodities are not available in the base year 0, in which case it is not possible simply to alter only the one aspect of the addition of commodities. The alteration must be a compound of the addition of new commodities, a change in the weighting system, and a shift in the base year. This second alternative will be discussed in a later section dealing with the three simultaneous changes.

There remains one circumstance where the present case might prove useful. This is where, even though the base year prices and values of the added commodities are available, and therefore the change involves only the addition of new commodities without the complication of an alteration in the weighting system and of a shift in the base year, the collection of the price data of the added commodities in the intervening years from year 1 to year $(T - 1)$ is either too difficult or too laborious. The ratio procedure is thus used even here to compile a continuous series from year

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

0 to year T and onwards. We need to study the efficiency or otherwise of the procedure in this case.

We examine the quantity

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 + \frac{\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}}{\frac{\sum_i {}_iP_0\,{}_iQ_0}{\sum_j {}_jP_0\,{}_jQ_0} + \frac{J_{t0}}{I_{t0}}}$$

Where the J's are again sub-indices referring to the added commodities and compiled by the method under discussion.

That is:

$$J_{T0} = \frac{1}{\sum {}_jP_0\,{}_jQ_0} \sum_j \frac{{}_jP_T}{{}_jP_0} \cdot {}_jP_0\,{}_jQ_0$$

and similarly for J_{t0}.

It will be seen that R in the present case is similar to that in the case of the unweighted arithmetic mean of relatives, except that in place of N/n, we have here $\sum_i {}_iP_0\,{}_iQ_0/\sum_j {}_jP_0\,{}_jQ_0$, and that the indices and sub-indices here refer to those compiled by weighting by base-year values. The same table as Table 1 can thus be used, provided that the above two points are borne in mind, and the same interpretation applies.

(ii) Weighted by arbitrary year values. All that has been said above in the case of weighting by base year values applies here, with one slight qualification. When the prices and values of the added commodities are not available for the arbitrary year a and the new index number therefore involves, besides the addition of the commodities, an alteration in the weighting system—say that of a more recent arbitrary year b—the base year 0 may or may not be changed. For base-weighting, a change in the weighting system automatically implies a shift in the base year.

All that this means is that the simultaneous changes may involve the two aspects of addition and weight-change, or three aspects with base-year-shift included. Both problems will be discussed later.

Where the weighting system (based on year a values) is not altered, the discussion and solution of the problem are exactly the same as those of weighting by base year values. Instead of the quantity $\sum_i {}_iP_0\,{}_iQ_0/\sum_j {}_jP_0\,{}_jQ_0$, we now have $\sum_i {}_iP_a\,{}_iQ_a/\sum_j {}_jP_a\,{}_jQ_a$.

(iii) Weighting by current year values. Here we have the following equality:

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 + \frac{\frac{\frac{J_{T0}}{I_{T0}} - 1}{\frac{\sum_i {}_iP_T\,{}_iQ_T}{\sum_j {}_jP_T\,{}_jQ_T} + 1} - \frac{\frac{J_{t0}}{I_{t0}} - 1}{\frac{\sum_i {}_iP_t\,{}_iQ_t}{\sum_j {}_jP_t\,{}_jQ_t} + 1}}{1 + \frac{\frac{J_{t0}}{I_{t0}} - 1}{\frac{\sum_i {}_iP_t\,{}_iQ_t}{\sum_j {}_jP_t\,{}_jQ_t} + 1}}$$

$$= 1 + E.$$

E, the error term, is affected by four quantities, viz., J_{T0}/I_{T0}, J_{t0}/I_{t0}, $\sum_i {}_iP_T\,{}_iQ_T/\sum_j {}_jP_T\,{}_jQ_T$ and $\sum_i {}_iP_t\,{}_iQ_t/\sum_j {}_jP_t\,{}_jQ_t$, as opposed to the earlier cases, where the equivalent of the last quantity did not appear.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We notice first of all that if the ratio of weights from year to year does not vary, i.e., if

$$\sum_i {}_iP_t\,{}_iQ_t/\sum_j {}_jP_t\,{}_jQ_t = \sum_i {}_iP_T\,{}_iQ_T/\sum_j {}_jP_T\,{}_jQ_T,$$

E would be the same as in the case of base year or arbitrary year weights. If, however, the weight ratio varies, the errors in estimation would also vary.

Table 3 shows, for variations in the values of

$$\frac{\sum_i {}_iP_T\,{}_iQ_T}{\sum_j {}_jP_T\,{}_jQ_T},\ \frac{\sum_i {}_iP_t\,{}_iQ_t}{\sum {}_jP_t\,{}_jQ_t} \text{ and } \frac{J_{T0}}{I_{T0}},$$

the limits of J_{t0}/I_{t0} for $\pm$ 5 per cent. error in R.

TABLE 3

Ranges of J_{t0}/I_{t0} *corresponding to:* (a) *Variations in the Values of the Three Ratios as Tabulated, and* (b) $\pm$ 5 *per cent. for the Error Term* E. *Addition of Commodities in Arithmetic Mean Index Number Weighted by Current Year Weights.* (*Figures on the top row denote the lower limits, and those on the bottom row the upper limits.* a *denotes non-existence of the particular positive values of* J_{t0}/I_{t0}.)

$\frac{\sum_i P_T\,{}_iQ_T}{\sum_j P_T\,{}_jQ_T} = 10$

	$\sum_i P_t\,{}_iQ_t/\sum_j P_t\,{}_jQ_t$				
J_{T0}/I_{T0}	6	8	10	12	14
0·5	0·4	0·2	a	a	a
	1·0	1·0	1·1	1·1	1·1
1·0	0·7	0·6	0·5	0·4	0·3
	1·4	1·5	1·6	1·7	1·8
3·0	1·9	2·1	2·4	2·6	2·9
	2·7	3·2	3·7	4·2	4·7
5·0	3·1	3·7	4·3	4·9	5·5
	4·0	4·9	5·8	6·7	7·5
10·0	6·1	7·6	9·1	10·5	12·0
	7·4	9·2	11·1	12·9	14·7

$\frac{\sum_i P_T\,{}_iQ_T}{\sum_j P_T\,{}_jQ_T} = 30$

	$\sum_i P_t\,{}_iQ_t/\sum_j P_t\,{}_jQ_t$				
J_{T0}/I_{T0}	26	28	30	32	34
0·5	a	a	a	a	a
	2·0	2·0	2·1	2·2	2·2
1·0	a	a	a	a	a
	2·4	2·5	2·6	2·7	2·8
3·0	1·4	1·4	1·4	1·5	1·5
	4·3	4·5	4·7	5·0	5·2
5·0	3·0	3·2	3·3	3·5	3·6
	6·1	6·5	6·8	7·2	7·6
10·0	7·2	7·6	7·9	8·6	9·0
	10·7	11·4	12·1	12·8	13·5

$\frac{\sum_i P_T\,{}_iQ_T}{\sum_j P_T\,{}_jQ_T} = 20$

	$\sum_i P_t\,{}_iQ_t/\sum_j P_t\,{}_jQ_t$				
J_{T0}/I_{T0}	16	18	20	22	24
0·5	a	a	a	a	a
	1·5	1·5	1·6	1·6	1·7
1·0	0·2	0·1	a	a	a
	1·9	2·0	2·1	2·2	2·3
3·0	1·7	1·8	1·9	2·0	2·1
	3·6	3·9	4·2	4·5	4·8
5·0	3·3	3·5	3·8	4·1	4·3
	5·3	5·8	6·3	6·8	7·3
10·0	7·1	7·9	8·6	9·3	10·0
	9·6	10·6	11·6	12·6	13·6

$\frac{\sum_i P_T\,{}_iQ_T}{\sum_j P_T\,{}_jQ_T} = 50$

	$\sum_i P_t\,{}_iQ_t/\sum_j P_t\,{}_jQ_t$				
J_{T0}/I_{T0}	46	48	50	52	54
0·5	a	a	a	a	a
	3·0	3·1	3·2	3·2	3·3
1·0	a	a	a	a	a
	3·5	3·6	3·7	3·8	3·9
3·0	0·5	0·5	0·5	0·5	0·4
	5·4	5·6	5·8	6·0	6·2
5·0	2·3	2·3	2·4	2·4	2·5
	7·4	7·6	7·9	8·2	8·4
10·0	6·7	6·9	7·0	7·4	7·6
	12·2	12·7	13·2	13·6	14·1

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$\text{For any value of } \frac{\sum_i {}_iP_T\,{}_iQ_T}{\sum_j {}_jP_T\,{}_jQ_T}, \quad \frac{\sum_i {}_iP_t\,{}_iQ_t}{\sum_j {}_jP_t\,{}_jQ_t}$$

is made to vary around it to only a moderate extent, as this probably conforms with reality.

Two points emerge from the table generally.

Firstly, when $\sum_i {}_iP_t\,{}_iQ_t/\sum_j {}_jP_t\,{}_jQ_t$ is less than $\sum {}_iP_T\,{}_iQ_T/\sum {}_jP_T\,{}_jQ_T$, the ranges of J_{t0}/I_{t0} are rather smaller than in the reverse situation.

Secondly, when $\sum {}_iP_t\,{}_iQ_t/\sum {}_jP_t\,{}_jQ_t$ is less than $\sum {}_iP_T\,{}_iQ_T/\sum {}_jP_T\,{}_jQ_T$, J_{t0}/I_{t0} tends to be smaller than when the reverse is the case, and in fact in some cases (as when $\sum {}_iP_T\,{}_iQ_T/\sum {}_jP_T\,{}_jQ_T = 10$ and 20), its upper limit is less than J_{T0}/I_{T0}.

It is generally true that the added commodities assume more and more importance with time. That is, t being any year before T, we can say that, generally,

$$\frac{\sum {}_jP_t\,{}_jQ_t}{\sum {}_iP_t\,{}_iQ_t} < \frac{\sum {}_jP_T\,{}_jQ_T}{\sum {}_iP_T\,{}_iQ_T} \quad \text{or} \quad \frac{\sum {}_iP_t\,{}_iQ_t}{\sum {}_jP_t\,{}_jQ_t} > \frac{\sum {}_iP_T\,{}_iQ_T}{\sum {}_jP_T\,{}_jQ_T}$$

This being so, we can conclude from the above discussion that : (a) the ranges of J_{t0}/I_{t0} tend to be larger for equivalent errors for current-weighted index numbers than for base-weighted or arbitrary year-weighted index numbers, and (b), on the other hand, J_{t0}/I_{t0} itself tends to be larger than J_{T0}/I_{T0} for certain values of the latter. For example, when we have $\sum {}_iP_T\,{}_iQ_T/\sum {}_jP_T\,{}_jQ_T = 10$, $\sum {}_iP_t\,{}_iQ_t/\sum {}_jP_t\,{}_jQ_t = 14$, $J_{T0}/I_{T0} = 5{\cdot}0$, J_{t0}/I_{t0} would need to be within the range of $5{\cdot}5$ and $7{\cdot}5$ for a maximum of 5 per cent. error to be tolerated. If, let us say, J_{t0}/I_{t0} were 4, then the error would be about 14 per cent.; or if J_{t0}/I_{t0} were also 5, the error would be about 8 per cent., while for base-weighted or arbitrary-year-weighted index numbers in the same circumstances the error would be about or less than 5 per cent.

(*b*) The geometric mean.

(i) Weighted by base year values. What has been said in connection with the arithmetic mean weighted by base year values applies here. Where continuity by use of the ratio method is required even though the base year values, and therefore prices, are known, we have

$$R = \left[\frac{\dfrac{J_{T0}}{I_{T0}}}{\dfrac{J_{t0}}{I_{t0}}}\right]^{\frac{1}{\frac{\sum_i {}_iP_0\,{}_iQ_0}{\sum_j {}_jP_0\,{}_jQ_0}+1}}$$

where the J's are, as before, sub-indices referring to the added commodities.

A comparison with the case of the unweighted geometric mean shows similarity between the values of R, except that here, instead of N/n, we have $\sum_i {}_iP_0\,{}_iQ_0/\sum_j {}_jP_0\,{}_jQ_0$. The same table (Table 2), with the corresponding qualification, is therefore applicable.

(ii) Weighted by arbitrary year values. Reference is made to the case of the arithmetic mean weighted by arbitrary year values, to which the present situation corresponds. Where even though the original arbitrary year (*a*) values and prices are available, the use of the ratio method for continuity is required, the position is exactly the same as the previous case of geometric mean weighted by base year values, with $\sum_i {}_iP_a\,{}_iQ_a/\sum_j {}_jP_a\,{}_jQ_a$ taking the place of $\sum_i {}_iP_0\,{}_iQ_0/\sum_j {}_jP_0\,{}_jQ_0$.

(iii) Weighted by current year values. It can be verified that in the present case,

$$R = \frac{\left(\dfrac{J_{T0}}{I_{T0}}\right)^{\frac{1}{\frac{\sum {}_iP_T\,{}_iQ_T}{\sum {}_jP_T\,{}_jQ_T}+1}}}{\left(\dfrac{J_{t0}}{I_{t0}}\right)^{\frac{1}{\frac{\sum {}_iP_t\,{}_iQ_t}{\sum {}_jP_t\,{}_jQ_t}+1}}}.$$

Here again we have four quantities affecting R, as in the case of the arithmetic mean.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Table 4 is compiled in the same way as Table 3; it shows, for ± 5 per cent. error in R, the limits of J_{t0}/I_{t0} corresponding to variations in the values of the other three ratios. It is clear from a comparison of the two tables that much the same conclusions can be drawn in the case of the geometric mean as in that of the arithmetic mean, namely, that when

$$\Sigma_j P_T {}_jQ_T / \Sigma_i P_T {}_iQ_T > \Sigma_j P_t {}_jQ_t / \Sigma_i P_t {}_iQ_t,$$

the method of continuity gives better estimates than in the case of fixed-weighted index numbers. On the other hand, when

$$\Sigma_j P_T {}_jQ_T / \Sigma_i P_T {}_iQ_T < \Sigma_j P_t {}_jQ_t / \Sigma_i P_t {}_iQ_t,$$

the estimate is less accurate.

TABLE 4

Ranges of J_{t0}/I_{t0} *corresponding to* (*a*) *Variations in the Values of the Three Ratios as Tabulated, and* (*b*) ± 5 *per cent. for the Error in Estimation. Addition of Commodities in the Geometric Mean Index Number Weighted by Current Year Weights.* (*Figures on the top row denote the lower limits, and those in the bottom row the upper limits.*)

$\frac{\Sigma_i P_T {}_iQ_T}{\Sigma_j P_T {}_jQ_t} = 10$

JT_0/IT_0	$\Sigma_i P_t\ Q_t/\Sigma_j P_t {}_jQ_t$ 6	8	10	12	14
0·5	0·45	0·36	0·29	0·23	0·19
	0·92	0·90	0·88	0·86	0·84
1·0	0·71	0·64	0·58	0·53	0·48
	1·43	1·58	1·76	1·95	2·16
3·0	1·43	1·58	1·76	1·94	2·15
	2·88	3·90	5·28	7·14	9·66
5·0	1·98	2·40	2·92	3·55	4·32
	3·99	5·92	8·80	13·06	19·40
10·0	3·08	4·24	5·85	8·06	11·11
	6·20	10·44	17·60	29·63	49·90

$\frac{\Sigma_i P_T {}_iQ_T}{\Sigma_j P_T {}_jQ_T} = 30$

JT_0/IT_0	$\Sigma_i P_t {}_iQ_t/\Sigma_j P_t {}_jQ_t$ 26	28	30	32	34
0·5	0·15	0·13	0·11	0·10	0·08
	2·19	2·32	2·46	2·61	2·76
1·0	0·27	0·24	0·22	0·20	0·18
	4·0	4·4	4·9	5·4	6·0
3·0	0·69	0·67	0·66	0·63	0·62
	10·4	12·4	14·7	17·5	20·8
5·0	1·1	1·1	1·1	1·1	1·1
	16·3	20·0	24·6	30·2	37·1
10·0	2·0	2·1	2·2	2·3	2·4
	29·7	38·2	49·1	63·2	81·2

$\frac{\Sigma_i P_T {}_iQ_T}{\Sigma_j P_T {}_jQ_T} = 20$

JT_0/IT_0	$\Sigma_i P_t {}_iQ_t/\Sigma_j P_t {}_jQ_t$ 16	18	20	22	24
0·5	0·25	0·21	0·18	0·15	0·13
	1·37	1·42	1·47	1·53	1·58
1·0	0·44	0·40	0·36	0·33	0·30
	2·4	2·7	2·9	3·3	3·6
3·0	1·06	1·07	1·08	1·08	1·09
	5·8	7·2	8·8	10·9	13·3
5·0	1·6	1·7	1·8	1·9	2·0
	8·8	11·4	14·7	19·0	24·5
10·0	2·8	3·2	3·6	4·1	4·6
	15·4	21·3	29·4	40·6	56·0

$\frac{\Sigma_i P_T {}_iQ_T}{\Sigma_j P_T {}_jQT} = 50$

JT_0/IT_0	$\Sigma_i P_t {}_iQ_t/\Sigma_j P_t {}_jQ_t$ 46	48	50	52	54
0·5	0·05	0·05	0·04	0·04	0·03
	5·9	6·4	6·9	7·4	8·0
1·0	0·10	0·09	0·08	0·08	0·07
	11·2	12·4	13·7	15·2	16·9
3·0	0·28	0·26	0·25	0·24	0·22
	30·7	35·5	41·2	47·6	55·0
5·0	0·44	0·43	0·41	0·40	0·39
	49·2	58·1	68·6	81·0	95·6
10·0	0·8	0·8	0·8	0·8	0·8
	93·3	113·2	137·2	166·4	201·8

c.—Ratio of aggregatives, using as basket of goods:

(i) base year quantities (Lespeyre Index Number). This index number is identical to that of weighted arithmetic mean using base year values as weights. The problem has already been discussed under that method.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(ii) arbitrary year quantities. This index number can be developed in form to be similar to the Lespeyre index number (and therefore to the base-value-weighted arithmetic mean index number), as follows:

$$I_{t0} = \frac{\sum_i {}_iP_t\ {}_iQ_a}{\sum_i {}_iP_0\ {}_iQ_a} = \frac{1}{\sum_i {}_iP_0\ {}_iQ_a} \sum_i \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_0\ {}_iQ_a.$$

Such a ratio-of-aggregative index number is therefore identical with an index number constructed by the weighted arithmetic mean using as weights the values ${}_iP_0\ {}_iQ_a$. While the discussion follows the line developed for the case of the weighted arithmetic mean using as weights the values ${}_iP_0\ {}_iQ_0$, this change in the weighting system has to be kept in mind. We have, for instance,

$$R = 1 + \frac{\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}}{\frac{\sum_i {}_iP_0\ {}_iQ_a}{\sum_j {}_jP_0\ {}_jQ_a} + \frac{J_{t0}}{I_{t0}}}$$

where the J's and the I's are sub-indices and indices relevant to the present method of construction. Further, in using Table 1, in place of N/n, we now have the quantity $\sum_i {}_iP_0\ {}_iQ_a / \sum_j {}_jP_0\ {}_jQ_a$.

(iii) Current year quantities (Paasche Index Number). In the same way that the last two index numbers can be derived to correspond with the weighted arithmetic mean index numbers using base year (or arbitrary year) values as weights, the present index number can be seen to correspond with the weighted arithmetic mean using current year values as weights. For the Paasche index number is, in fact, a weighted arithmetic mean index having as weights values such as $P_0\ Q_T$ or $P_0\ Q_t$ as the case may be.

3. *Deletion of Commodities*

i will again be used to denote the original list of commodities, $i = 1, 2, \ldots, N$. p of these N commodities are to be deleted, and they will be denoted by k, $k = 1, 2, \ldots, p$, i.e., it is assumed that the first n of the original commodities are affected. The revised list now consists of $(N - p)$ commodities, and is indicated by i', $i' = p + 1, p + 2, \ldots, N$.

Throughout this section, we shall use the symbol K to denote the sub-index based on the deleted commodities. Wherever K occurs, it is understood that it is constructed by the method under discussion. Thus, if the method of construction under discussion is the arithmetic mean weighted by base year values, then

$$K_{T0}\ (\text{say}) = \frac{1}{\sum_k {}_kP_0\ {}_kQ_0} \sum_k \frac{{}_kP_T}{{}_kP_0} \cdot {}_kP_0\ {}_kQ_0.$$

and so on.

A.—The unweighted mean of relatives.

(a) The arithmetic mean. Here we have the equality:

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 - \frac{\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}}{\frac{N}{p} - \frac{K_{t0}}{I_{t0}}}.$$

Table 5 is drawn up in the same way as Table 1 to show the values of K_{t0}/I_{t0} that would yield $\pm$ 5 per cent. error in R for corresponding values of K_{T0}/I_{T0} and N/p. In the table the top figure in each case denotes minimum value yielding an error of $-$ 5 per cent., and the bottom figure denotes maximum value yielding an error of $+$ 5 per cent. error.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The difference between this table and Table 1 is that here values of K_{t0}/I_{t0} less than K_{T0}/I_{T0} give rise to under-estimation, and values of K_{t0}/I_{t0} larger than K_{T0}/I_{T0} give rise to over-estimation; the converse is the case for addition of commodities as shown in Table 1. Another point is that the ranges of values of K_{t0}/I_{t0} are slightly narrower for the same error-limits than those in the case of addition of commodities. This means that the error tends to be slightly larger in absolute terms for deletion of commodities than for addition of commodities. For example, in dealing with deletion of commodities, if $K_{T0}/I_{T0} = 2$, $K_{t0}/I_{t0} = 3$, and $N/p = 30$, the error in absolute terms is about 3·7 per cent. Correspondingly, in the case of addition of commodities, if $J_{T0}/I_{T0} = 2$, $J_{t0}/I_{t0} = 3$, and $N/n = 30$, the error in absolute terms is just about 3 per cent.

TABLE 5

Limits of K_{t0}/I_{t0} *corresponding to Some Values of* K_{T0}/I_{T0} *and of* N/p *for* E *to Fall within the Range* ± 5 *per cent. (Figures on the top row denote minimum values [corresponding to* − 5 *per cent. error] and figures on the bottom row maximum values [corresponding to* + 5 *per cent. error].)*

	Ranges of K_{t0}/I_{t0} *for Corresponding* K_{T0}/I_{T0} *and for* N/p *as follows:*									
K_{T0}/I_{T0}	10	20	30	40	50	60	70	80	90	100
0·1	a	a	a	a	a	a	a	a	a	a
	0·57	1·05	1·52	2·00	2·48	2·95	3·43	3·90	4·38	4·86
0·25	a	a	a	a	a	a	a	a	a	a
	0·71	1·20	1·67	2·14	2·62	3·10	3·57	4·05	4·52	5·00
0·5	a	a	a	a	a	a	a	a	a	a
	0·95	1·43	1·90	2·38	2·86	3·33	3·80	4·30	4·76	5·24
1·0	0·53	a	a	a	a	a	a	a	a	a
	1·43	1·90	2·38	2·86	3·33	3·80	4·30	4·76	5·24	5·71
1·5	1·05	0·53	a	a	a	a	a	a	a	a
	1·90	2·38	2·86	3·33	3·80	4·30	4·76	5·24	5·71	6·19
2·0	1·58	1·05	0·53	a	a	a	a	a	a	a
	2·38	2·86	3·33	3·80	4·30	4·76	5·24	5·71	6·19	6·67
2·5	2·10	1·58	1·05	0·53	a	a	a	a	a	a
	2·86	3·33	3·80	4·30	4·76	5·24	5·71	6·19	6·67	7·14
3·0	2·63	2·10	1·58	1·05	0·53	a	a	a	a	a
	3·33	3·80	4·30	4·76	5·24	5·71	6·19	6·67	7·14	7·62
3·5	3·16	2·63	2·10	1·58	1·05	0·53	a	0·53	a	a
	3·80	4·30	4·76	5·24	5·71	6·19	6·67	7·14	7·62	8·10
4·0	3·68	3·16	2·63	2·10	1·58	1·05	0·53	a	a	a
	4·30	4·76	5·24	5·71	6·19	6·67	7·14	7·62	8·10	8·57
5·0	4·74	4·21	3·68	3·16	2·63	2·10	1·58	1·05	a	a
	5·24	5·71	6·19	6·67	7·14	7·62	8·10	8·57	9·05	9·52
10·0	10·0	9·47	8·95	8·42	7·89	7·37	6·84	6·32	5·79	5·26
	10·0	10·48	10·95	11·43	11·90	12·38	12·86	13·33	13·81	14·29

(a) shows that there is no minimum positive value of K_t/I_t for corresponding K_T/I_T and N/p.

(*b*) The geometric mean. Here we have

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = \left[\frac{\frac{K_{t0}}{I_{t0}}}{\frac{K_{T0}}{I_{T0}}}\right]^{\frac{1}{\frac{N}{p}-1}}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

From this relationship Table 6 is constructed in the same way as the previous tables. It is to be compared directly with Table 2; from this comparison we obtain the same general conclusions as from the comparison of Table 3 with Table 1.

Moreover, in both situations of commodity changes studied so far, the geometric mean index number is found to yield a smaller error than the arithmetic mean index number. Corresponding to the example discussed under the arithmetic mean (i.e. where K_{T0}/I_{T0} or $J_{T0}/I_{T0} = 2$, K_{t0}/I_{t0} or $J_{t0}/I_{t0} = 3$ and $N/p = 30$) where the error was found to be about 3 to 4 per cent., for the geometric mean the error is something like 1·3 to 1·4 per cent.

B.—The weighted mean of relatives.

(*a*) The arithmetic mean:

(i) weighted by base year values. Here we have the quantity

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 - \frac{\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}}{\frac{\sum_i {}_iP_0\,{}_iQ_0}{\sum_k {}_kP_0\,{}_kQ_0} - \frac{K_{t0}}{I_{t0}}}.$$

and (ii) weighted by arbitrary year values. Here

$$R = 1 - \frac{\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}}{\frac{\sum_i {}_iP_a\,{}_iQ_a}{\sum_k {}_kP_a\,{}_kQ_a} - \frac{K_{t0}}{I_{t0}}}.$$

In both cases the assumption is that the base year or the arbitrary year does not become involved in the change. That is, we need not concern ourselves at this stage with simultaneous changes.

In both cases again, we see that Table 5 is applicable except that in place of N/p we substitute either $\sum_i {}_iP_0\,{}_iQ_0/\sum_k {}_kP_0\,{}_kQ_0$ or $\sum_i {}_iP_a\,{}_iQ_a/\sum_k {}_kP_a\,{}_kQ_a$. The general conclusions are similar to those derived for the case of the unweighted arithmetic mean, and comparison is to be made with the corresponding methods in the case of commodity addition.

(iii) Weighted by current year weights. Corresponding to the same situation under the addition of commodities, we study the ratio

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{T0}}} = 1 - \frac{\frac{\frac{K_{T0}}{I_{T0}} - 1}{\frac{\sum_i {}_iP_T\,{}_iQ_T}{\sum_k {}_kP_T\,{}_kQ_T} - 1} - \frac{\frac{K_{t0}}{I_{t0}} - 1}{\frac{\sum_i {}_iP_t\,{}_iQ_t}{\sum_k {}_kP_t\,{}_kQ_t} - 1}}{1 - \frac{\frac{K_{t0}}{I_{t0}} - 1}{\frac{\sum_i {}_iP_t\,{}_iQ}{\sum_k {}_kP_t\,{}_kQ_t} - 1}}$$

$$= 1 - E.$$

The limits of J_{t0}/I_{t0} for $\pm$ 5 per cent. for the error term E and corresponding to variations in the values of the other three terms are tabulated in Table 7, which is to be compared with Tables 3 and 5. It can be seen from this comparison that more or less the same conclusions can be drawn as in the case of the addition of commodities, except that the tendency when $\Sigma_k {}_kP_t\,{}_kQ_t/\Sigma_k {}_kP_t\,{}_kQ_t$ as in the case of the addition of commodities, except that the tendency when $\Sigma_k {}_kP_t\,{}_kQ_t/\Sigma_k {}_kP_t\,{}_kQ_t$ is greater than $\Sigma_i {}_iP_T\,{}_iQ_T/\Sigma_k {}_kP_T\,{}_kQ_T$ as compared to that in the reverse situation is not so large as to be obvious without a fairly close scrutiny. This conforms with what we have already seen, that in deletion of commodities the error tends to be slightly larger than in addition of commodities.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 6

Limits of K_{t0}/I_{t0} *corresponding to Some Values of* K_{T0}/I_{T0} *and of* N/p *for* E *to fall within the Range* ± 5 *per cent. (Figures on the first row denote minimum values [corresponding to* − 5 *per cent.]; figures on the second row denote maximum values [corresponding to* + 5 *per cent.].)*

Ranges of K_{t0}/I_{t0} *for Corresponding* K_{T0}/I_{T0} *and for* N/p *as follows:*

K_{T0}/I_{T0}	10	20	30	50	100
0·1	0·063	0·038	0·023	0·008	0·001
	0·155	0·253	0·412	1·093	12·55
0·5	0·315	0·189	0·113	0·04	0·003
	0·776	1·264	2·06	5·47	62·77
1·0	0·63	0·38	0·23	0·08	0·006
	1·55	2·53	4·12	10·93	125·5
2·0	1·26	0·75	0·45	0·16	0·012
	3·10	5·06	8·24	21·87	251·1
3·0	1·89	1·13	0·68	0·24	0·019
	4·66	7·58	12·36	32·80	376·6
4·0	2·51	1·51	0·90	0·32	0·025
	6·21	10·11	16·47	43·74	502·3
5·0	3·15	1·89	1·12	0·40	0·031
	7·76	12·64	20·60	54·68	627·7
10·0	6·30	3·77	2·26	0·81	0·062
	15·52	25·28	41·19	109·30	1255·0

(*b*) The geometric mean.

Here again the base year weighting and arbitrary year weighting are similar to the method of unweighted geometric mean. For we have:

(i) weighted by base year values,

$$R = \left[\frac{\dfrac{K_{t0}}{I_{t0}}}{\dfrac{K_{T0}}{I_{T0}}}\right]^{\frac{1}{\sum_i {}_iP_0\,{}_iQ_0}{\sum_k {}_kP_0\,{}_kQ_0}-1}.$$

and (ii) weighted by arbitrary year values,

$$R = \left[\frac{\dfrac{K_{t0}}{I_{t0}}}{\dfrac{K_{T0}}{I_{T0}}}\right]^{\frac{1}{\sum_i {}_iP_a\,{}_iQ_a}{\sum_k {}_kP_a\,{}_kQ_a}-1}.$$

Table 6 is therefore to be used in connection with both these two situations, with $\sum_i {}_iP_0\,{}_iQ_0/\sum_k {}_kP_0\,{}_kQ_0$ and $\sum_i {}_iP_a\,{}_iQ_a/\sum_k {}_kP_a\,{}_kQ_a$ taking the place of N/p.

(iii) Weighted by current year values. Again, corresponding to similar cases already discussed, we study the relationship

$$R = \left(\frac{K_{t0}}{I_{t0}}\right)^{\frac{1}{\sum_i {}_iP_t\,{}_iQ_t}{\sum_k {}_kP_t\,{}_kQ_t}-1} \Bigg/ \left(\frac{K_{T0}}{I_{T0}}\right)^{\frac{1}{\sum_i {}_iP_T\,{}_iQ_T}{\sum_k {}_kP_T\,{}_kQ_T}-1}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

from which Table 8 has been compiled to show the limits of K_{t0}/I_{t0} for $\pm$ 5 per cent. error in estimation and for variations in the other three ratios. Here although the limits are also slightly narrower throughout than in the corresponding situation in the addition of commodities, they are still very broad. Further, a comparison of the two situations shows that practically the same conclusions apply here as have been reached in respect of Table 4.

TABLE 7

Ranges of K_{t0}/I_{t0} *corresponding to:* (a) *Variations in the Values of the Three Ratios as Tabulated, and* (b) $\pm$ 5 *per cent. for the Error Term* E. *Deletion of Commodities in Arithmetic Mean Index Number Weighted by Current Year Weights.* (*Figures on the top row denote the lower limits, and those on the bottom row upper limits.*)

$\frac{\Sigma_i P_{T\,i} Q_T}{\Sigma_k P_{T\,k} Q_T} = 10$

	$\Sigma_i P_{t\,i} Q_t / \Sigma_k P_{t\,k} Q_t$				
K_{T_0}/I_{T_0}	6	8	10	12	14
0·5	0·4	0·2	a	a	a
	1·0	0·96	0·95	0·94	0·93
1·0	0·7	0·6	0·5	0·4	0·3
	1·2	1·3	1·4	1·5	1·6
3·0	1·9	2·3	2·6	3·0	3·4
	2·3	2·8	3·3	3·9	4·4
5·0	3·1	3·9	4·7	5·6	6·4
	3·4	4·3	5·2	6·2	7·1
10·0	6·0	8·0	10·0	12·0	14·0
	6·0	8·0	10·0	12·0	14·0

$\frac{\Sigma_i P_{T\,i} Q_T}{\Sigma_k P_{T\,k} Q_T} = 30$

	$\Sigma_i P_{t\,i} Q_t / \Sigma_k P_{t\,k} Q_t$				
K_{T_0}/I_{T_0}	26	28	30	32	34
0·5	a	a	a	a	a
	1·8	1·8	1·9	2·0	2·0
1·0	a	a	a	a	a
	2·2	2·3	2·4	2·5	2·6
3·0	1·5	1·5	1·6	1·6	1·7
	3·8	4·1	4·3	4·5	4·7
5·0	3·3	3·5	3·7	3·9	4·1
	5·5	5·8	6·2	6·5	6·9
10·0	7·9	8·4	8·9	9·5	10·0
	9·6	10·3	11·0	11·6	12·3

$\frac{\Sigma_i P_{T\,i} Q_T}{\Sigma_j P_{T\,j} Q_T} = 20$

	$\Sigma_i P_{t\,i} Q_t / \Sigma_j P_{t\,j} Q_t$				
J_{T_0}/I_{T_0}	16	18	20	22	24
0·5	a	a	a	a	a
	1·3	1·4	1·4	1·5	1·5
1·0	0·2	0·1	a	a	a
	1·7	1·8	1·9	2·0	2·1
3·0	1·9	2·0	2·1	2·2	2·3
	3·2	3·5	3·8	4·1	4·4
5·0	3·5	3·9	4·2	4·5	4·9
	4·7	5·2	5·7	6·2	6·7
10·0	7·7	8·6	9·5	10·4	11·3
	8·5	9·5	10·5	11·5	12·5

$\frac{\Sigma_i P_{T\,i} Q_T}{\Sigma_j P_{T\,j} Q_T} = 50$

	$\Sigma_i P_{t\,i} Q_t / \Sigma_j P_{t\,j} Q_t$				
J_{T_0}/I_{T_0}	46	48	50	52	54
0·5	a	a	a	a	a
	2·7	2·8	2·9	2·9	2·9
1·0	a	a	a	a	a
	3·1	3·2	3·3	3·4	3·4
3·0	0·6	0·5	0·5	0·5	0·5
	4·9	5·1	5·2	5·4	5·4
5·0	2·5	2·6	2·6	2·7	2·8
	6·6	6·9	7·1	7·4	7·4
10·0	7·3	7·6	7·9	8·2	8·5
	11·0	11·5	11·9	12·3	12·3

(a) denotes non-existence of the particular positive values of J_{t0}/I_{t0}.

C.—Ratio of aggregatives.

All that has been said in the case of addition of commodities regarding the correspondence of this type of index numbers with the weighted arithmetic mean index numbers applies in the present case and indeed throughout this article.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 8

Ranges of K_{t0}/I_{t0} *corresponding to:* (a) *Variations in the Values of the Three Ratios as Tabulated, and* (b) $\pm$ 5 *per cent. for the Error in Estimation. Deletion of Commodities in the Geometric Mean Index Number Weighted by Current Year Weights. (Figures in the top row denote lower limits, and those in the bottom row upper limits.)*

$\frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_k P_T\, {}_kQ_T} = 10$

	$\Sigma_i P_t\, {}_iQ_t/\Sigma_k P_t\, {}_kQ_t$				
K_{T0}/I_{T0}	6	8	10	12	14
0·5	0·5	0·4	0·3	0·2	0·2
	0·9	0·8	0·8	0·7	0·7
1·0	0·8	0·7	0·6	0·6	0·5
	1·3	1·4	1·6	1·7	1·9
3·0	1·4	1·6	1·9	2·2	2·5
	2·4	3·3	4·7	6·6	9·2
5·0	1·9	2·4	3·2	4·1	5·2
	3·1	4·9	7·8	12·2	19·3
10·0	2·8	4·2	6·3	9·5	14·3
	4·6	8·4	15·5	28·5	52·5

$\frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_k P_T\, {}_kQ_T} = 30$

	$\Sigma_i P_t\, {}_iQ_t/\Sigma_k P_t\, {}_kQ_t$				
K_{T0}/I_{T0}	26	28	30	32	34
0·5	0·15	0·13	0·11	0·10	0·08
	1·9	2·0	2·1	2·2	2·3
1·0	0·3	0·3	0·2	0·2	0·2
	3·4	3·7	4·1	4·5	5·0
3·0	0·7	0·7	0·7	0·7	0·6
	8·7	10·4	12·4	14·7	17·5
5·0	1·1	1·1	1·1	1·1	1·1
	13·6	16·7	20·6	25·4	31·3
10·0	2·0	2·1	2·3	2·4	2·5
	24·7	31·9	41·2	53·2	68·8

$\frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_j P_T\, {}_jQ_T} = 20$

	$\Sigma_i P_t\, {}_iQ_t/\Sigma_j P_t\, {}_jQ_t$				
J_{T0}/I_{T0}	16	18	20	22	24
0·5	0·3	0·2	0·2	0·2	0·1
	1·2	1·2	1·3	1·3	1·3
1·0	0·5	0·4	0·4	0·3	0·3
	2·1	2·3	2·5	2·8	3·1
3·0	1·1	1·1	1·1	1·1	1·2
	5·0	6·1	7·6	9·4	11·6
5·0	1·6	1·8	1·9	2·0	2·2
	7·4	9·7	12·6	16·5	21·6
10·0	2·9	3·3	3·8	4·3	5·0
	12·8	18·0	25·3	35·5	49·9

$\frac{\Sigma_i P_T\, {}_iQ_T}{\Sigma_j P_T\, {}_jQ_T} = 50$

	$\Sigma_i P_t\, {}_iQ_t/\Sigma_j P_t\, {}_jQ_t$				
J_{T0}/I_{T0}	46	48	50	52	54
0·5	0·05	0·05	0·04	0·04	0·03
	4·8	5·1	5·5	5·9	6·3
1·0	0·10	0·09	0·08	0·07	0·07
	9·0	9·9	10·9	12·1	13·3
3·0	0·27	0·26	0·24	0·23	0·22
	24·7	28·4	32·8	37·8	43·6
5·0	0·43	0·42	0·40	0·39	0·38
	39·4	46·4	5·47	64·3	75·8
10·0	0·8	0·8	0·8	0·8	0·8
	74·5	90·2	109·3	132·4	160·3

4. *Change in the Weighting System*

Only index numbers involving arbitrary year weighting are concerned here. In the base year weighting, a change in the weighting system is simultaneous with a change in the base year.

A.—Weighted mean of relatives.

The commodity i has as its original weight ${}_iP_a\, {}_iQ_a$ referring to arbitrary year a. In the new index number, this weight is changed to ${}_iP_b\, {}_iQ_b$ referring to a different year b. For the purpose of our discussion we introduce the idea of percentage weighting, by having

$$ {}_iW_a = \frac{{}_iP_a\, {}_iQ_a}{\Sigma_i P_a\, {}_iQ_a} \cdot 100, \quad {}_iW_b = \frac{{}_iP_b\, {}_iQ_b}{\Sigma_i P_b\, {}_iQ_b} \cdot 100, $$

and so on. We also define ${}_iv = {}_iW_b - {}_iW_a$, so that $\Sigma_i v = 0$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(*a*) The arithmetic mean. Here the following relationship obtains:

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 + \frac{\Sigma\, {}_iv \left(\frac{{}_iP_T}{{}_iP_0} \Big/ I_{T0} - \frac{{}_iP_t}{{}_iP_0} \Big/ I_{t0} \right)}{1 + \Sigma\, {}_iv \left(\frac{{}_iP_t}{{}_iP_0} \Big/ I_{t0} \right)}.$$

$$= 1 + E.$$

We now define

$$\frac{{}_iP_T}{\frac{{}_iP_0}{I_{T0}}} = {}_i\rho_T; \qquad \frac{{}_iP_t}{\frac{{}_iP_0}{I_{t0}}} = {}_i\rho_t.$$

Then E can be expressed as follows:

$$E = \frac{\sigma_{\rho_T} r_{v\rho_T} - \sigma_{\rho_t} r_{v\rho_t}}{\frac{1}{\sigma v} + \sigma_{\rho_t} r_{v\rho_t}},$$

This error term involves several unknown quantities, and to assess its approximate magnitude under various situations we have to introduce the following two assumptions: (a) firstly, the assumption, perhaps not too unreasonable, that $\sigma_{\rho_T} \doteqdot \sigma_{\rho_t}$, i.e., that the proportionate price-ratios vary more or less to the same extent from year to year; (b) next, to assess σ_{ρ_t}, we see that the numerator of ${}_i\rho_t$, namely ${}_iP_t/{}_iP_0$ can vary fairly extensively say from 0·5 to 10·0, and that the denominator, I_{t0}, is probably of order of magnitude of hundreds. The widest range of variability of ${}_i\rho_t$ is likely to be between say 0·001 and 0·020. If we now regard ρ_t as samples from a normal population, we can apply the findings of Tippett (1925) to estimate the value of σ_{ρ_t} from the range. Thus, if say $N = 10$, the ratio range/standard deviation = 3·078, from which we can estimate that σ_{ρ_t} is something like 0·006. If $N = 100$, range/σ = 5, so that σ_{ρ_t} is approximately 0·004. We shall discuss the present situation on the assumption that $N = 10$; if N is larger, it is easy to see that E would be smaller.

We obtain, as a result of the assumptions that (a) $N = 10$; (b) ρ_t varies from about 0·001 to 0·020, and (c) $\sigma_{\rho T} \doteqdot \sigma_{\rho t}$, the following form for E:

$$E = \frac{r_{v\varrho_T} - r_{v\varrho_t}}{\frac{1}{0{\cdot}006\ \sigma_v} + r_{v\rho_t}}.$$

In this relationship, σ_v is the quantity we are interested in. If say $\sigma_v = 10$, we can tolerate a variability of v to the extent of 10 per cent.

In Table 9 are shown the values for $r_{v\rho_t}$ that would give rise to $\pm$ 5 per cent. error corresponding to certain values of σ_v and of $r_{v\rho_T}$. The possible range goes from the lowest possible value of -1 to the highest possible value of $+1$. In the table, the top row in each cell denotes the lowest value corresponding to the situation under consideration, and the bottom row denotes the highest value. Where neither is stated, both the lowest possible value of -1 and the highest possible value of $+1$ would yield error of less than 5 per cent. Where both are not stated, it means that all possible values of $r_{v\rho_t}$ give rise to under 5 per cent. error. Thus, in part (a) of the table, when $r_{v\rho_T} = -0{\cdot}6$ and $\sigma_v = 20$, the limits of $r_{v\rho_t}$ are nearly -1 to about $-0{\cdot}2$. If $r_{v\rho_t}$ is positive, then the error would be larger than 5 per cent. in this situation. If $r_{v\rho_T} = -0{\cdot}4$ and $\sigma_v = 10$, the limits of $r_{v\rho_t}$ are between -1 and 0·46. When $\sigma_v = 5$, most of the limits of $r_{v\rho_t}$ stretch over nearly the whole possible range, and when $\sigma_v = 4$, all values are over the whole possible range, irrespective of the value of $r_{v\rho_T}$. This means that for σ_v of magnitude of about 4, practically any value of $r_{v\rho_T}$ and of $r_{v\rho_t}$ would give an error of under 5 per cent. For higher values of σ_v up to 10, except when $r_{v\rho_T}$ is very high positively or negatively, the same findings seem to apply. For still higher values of σ_v more care in application would have to be exercised.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE 9

Limits of $r_{v\rho_t}$ *for corresponding Values of* σ_v *and of* $r_{v\rho_T}$ *for Errors not Exceeding* $\pm$ 5 *per cent. in Ratio Continuity Conversion in the case of Weight change. Reference is made to the text for explanation and interpretation.*

	Values of $r_{v\rho_t}$ *for Corresponding* σ_v *and* $r_{v\rho_T}$ *as follows:*										
σ_v	−1·0	−0·8	−0·6	−0·4	−0·2	0	0·2	0·4	0·6	0·8	1·0
	(a) $N = 10$; range of price ratios from 0·001 to 0·020. Hence $\sigma_{p_T} \doteqdot \sigma_{p_t} = 0·006$										
50	—	−0·92	−0·73	−0·54	−0·35	−0·16	0·03	0·22	0·41	0·60	0·79
	−0·88	−0·67	−0·46	−0·25	−0·04	0·18	0·39	0·60	0·81	—	—
20	—	—	−0·97	−0·78	−0·59	−0·40	−0·21	−0·02	0·17	0·37	0·56
	−0·61	−0·40	−0·19	0·02	0·23	0·44	0·65	0·86	—	—	—
10	—	—	—	—	−0·98	−0·79	−0·60	−0·41	−0·22	−0·03	0·16
	−0·17	0·04	0·26	0·46	0·67	0·88	—	—	—	—	—
7·5	—	—	—	—	—	—	−0·87	−0·68	−0·49	−0·30	−0·11
	0·17	0·33	0·54	0·75	0·96	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—	—	−0·83	−0·63
	0·70	0·91	—	—	—	—	—	—	—	—	—
4	—	—	—	—	—	—	—	—	—	—	−1·00
	(b) $N = 50$; range of price ratios from 0·001 to 0·020. Hence $\sigma_{\rho_T} \doteqdot \sigma_{\rho_t} = 0·004$										
50	—	−1·00	−0·81	−0·62	−0·43	−0·24	−0·05	0·14	0·33	0·52	0·71
	−0·79	−0·58	−0·37	−0·16	0·05	0·26	0·47	0·68	0·89	—	—
20	—	—	—	−0·98	−0·79	−0·60	−0·40	−0·21	−0·02	0·17	0·36
	−0·39	−0·18	0·03	0·24	0·45	0·66	0·87	—	—	—	—
10	—	—	—	—	—	—	−1·00	−0·81	−0·62	−0·43	−0·24
	0·26	0·47	0·68	0·89	—	—	—	—	—	—	—
6·5	—	—	—	—	—	—	—	—	—	—	−0·88
	0·97	—	—	—	—	—	—	—	—	—	—
	(c) $N = 10$; range of price ratios 0·005 to 0·020. Hence $\sigma_{\rho_T} \doteqdot \sigma_{p_t} = 0·005$										
50	—	−0·95	−0·76	−0·57	−0·38	−0·19	0·00	0·19	0·38	0·57	0·76
	−0·84	−0·63	−0·42	−0·21	0·00	0·21	0·42	0·63	0·84	—	—
20	—	—	—	−0·86	−0·67	−0·48	−0·29	−0·10	0·10	0·29	0·48
	−0·53	−0·32	−0·11	0·11	0·32	0·53	0·74	0·95	—	—	—
10	—	—	—	—	—	−0·95	−0·76	−0·57	−0·38	−0·19	0·00
	0·00	0·21	0·42	0·63	0·84	—	—	—	—	—	—
7·5	—	—	—	—	—	—	—	−0·89	−0·70	−0·51	−0·32
	0·35	0·56	0·77	0·98	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—	—	—	−0·95
	(d) $N = 50$; range of price ratios from 0·005 to 0·020. Hence $\sigma_{\rho_T} \doteqdot \rho_t = 0·003$.										
50	—	—	−0·86	−0·67	−0·48	−0·29	−0·10	0·10	0·29	0·48	0·67
	−0·74	−0·53	−0·32	−0·11	0·11	0·32	0·53	0·74	0·95	—	—
20	—	—	—	—	−0·90	−0·71	−0·52	−0·33	−0·14	0·05	0·24
	−0·26	−0·05	0·16	0·37	0·58	0·79	1·00	—	—	—	—
10	—	—	—	—	—	—	—	—	−0·86	−0·67	−0·48
	0·53	0·74	0·95	—	—	—	—	—	—	—	—
8	—	—	—	—	—	—	—	—	—	—	−0·83
	0·92	—	—	—	—	—	—	—	—	—	—

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The above discussion is based on the assumption that $N = 10$, that the limits of ${}_iP_T/{}_iP_0$ and of ${}_iP_t/{}_iP_0$ are between 0·5 and 10·0; and that ${}_i\rho_T \doteqdot {}_i\rho_t$. If N is larger or if the price ratios range within narrower limits, the error would be smaller. The example where $N = 50$ is seen in part (b) of Table 9. Here for values of σ_v from about 10 downwards, except for very high positive or negative values of $r_{v\rho_T}$ (then only for σ_v near to 10), we need not concern ourselves with values of $r_{v\rho_T}$ or of $r_{v\rho_t}$. Part (c) of the table shows the effect of narrowing the range of price-ratio variations from between 0·001 and 0·020 to between 0·005 and 0·020, and part (d) shows the effect both of increasing N and of narrowing the range of price-ratio variations.

In general, it can be concluded that where a fair number of commodities are involved and the price-ratios do not vary too violently, weight changes to an extent that their standard deviation is about 10 per cent. would give rise to quite small error when using the ratio conversion method of continuity. Where weight changes are of larger magnitudes, the correlation between these changes and ρ might have to be taken into consideration. If these correlations are high either positively or negatively, the error tends to increase, if the two correlations (viz., $r_{v\rho_T}$ and $r_{v\rho_t}$) are of opposite signs. As an illustration, if under (a) above, we have $\sigma_v = 20$, and if $r_{v\rho_T} = -1$ and $r_{v\rho_t} = 1$, then the error is -21 per cent.; if $r_{v\rho_T} = -0{\cdot}5$ and $r_{v\rho_t} = 0{\cdot}5$, E is -11 per cent.

(b) The geometric mean. Here we can derive the following relationship:

$$R = \frac{\dfrac{I'_{T0}}{I_{T0}}}{\dfrac{I'_{t0}}{I_{t0}}} = \sqrt[100]{\prod_i \left(\frac{{}_iP_T}{{}_iP_t}\right)^{{}_iv}}$$

This factor depends to a large extent again on the correlation between the weight change and the price trend. If the weight change involves added weight for commodities which have been increasing in price, and reduced weight for those which have been decreasing in price (or vice versa in the case of negative correlation), the error will be larger than if the opposite were the case.

To assess the extent of this error, we shall consider a case where price changes and weight changes are small. We define

$${}_i\rho = \frac{{}_iP_T - {}_iP_t}{{}_iP_t}$$

so that

$$\frac{{}_iP_T}{{}_iP_t} = 1 + {}_i\rho.$$

${}_i\rho$ is assumed to be small; also if ${}_iP_T > {}_iP_t$, ${}_i\rho$ will be positive, and if ${}_iP_T < {}_iP_t$, it will be negative.

In this circumstance, then

$$\log_e R = \frac{1}{100}\left[\Sigma\, {}_iv \log_e(1 + {}_i\rho)\right].$$

$$= \frac{1}{100}\left[\Sigma\, {}_iv \left({}_i\rho - \frac{{}_i\rho^2}{2} + \frac{{}_i\rho^3}{3} \quad \cdots \quad \right)\right].$$

Since ${}_i\rho$ is small, we can ignore terms involving ${}_i\rho^2$ and higher powers of ρ. We have

$$\log_e R \doteqdot \frac{1}{100}\Sigma\, {}_iv\, {}_i\rho = \frac{1}{100}(N\, \sigma_v\, \sigma_\rho\, r_{v\rho}).$$

Assuming that N consists of 10 commodity items, that ${}_iv$ ranges from say -15 per cent. to $+15$ per cent., so that σ_v is approximately 9 or 10, and that ${}_i\rho$ ranges from about $-0{\cdot}1$ to $+0{\cdot}1$, so that σ_ρ is of the order of 0·07, we obtain

$$\log_e R \doteqdot \frac{6{\cdot}5\, r_{v\rho}}{100}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

or

$$\log_{10} R \doteqdot \frac{2 \cdot 8\, r_{vp}}{100}.$$

If r is large, say near 1, there would be an error of about 6 per cent. to 7 per cent.; if r is say 0·5, the error would be just over 3 per cent. The maximum error for this situation is therefore about 7 per cent.

In general, of course,

$$\log_{10} R = \frac{N\, \sigma_v\, \sigma_{\log p}\, r_{v \log p}}{100}$$

where ${}_ip$ here denotes ${}_iP_T/{}_iP_t$. If we postulate that $N = 10$, p varies from 1 to 4, and v from -15 to $+15$, then

$$\log_{10} R \doteqdot \frac{20 r_{v \log p}}{100}$$

with a maximum error of 59 per cent. $r_{v \log p}$ has to be fairly low for the error to be small; say $r = 0 \cdot 1$, then the error is about 5 per cent.

If v varies from -5 to $+5$, or if p varies from 1 to 1·58, then the maximum error will be something like 16 per cent. With $r_v \log_p = 0 \cdot 5$ the error is about 8 per cent.

Another point to note is that if N increases, the error also increases fairly rapidly. Thus, if $N = 10$, ${}_ip$ varies from 1 to 1·58, and v from -5 to $+5$, the maximum error is around 5 per cent.; but if $N = 50$, then the maximum error is of the order of 10·5 per cent.

We can then conclude that changing the weighting system for the geometric mean weighted by arbitrary year weights, if the number of commodity items is large, or if the price trend has been rather violent, or if the weight changes are fairly extensive, the ratio process is likely to lead to considerable errors unless the correlation between the weight change and the price-ratio is very small.

Even for the arithmetic mean, it was found that the size of the weight-changes can affect the error as well as the variability of the price ratios. But there we also found that the number of commodity items tended to decrease the error.

As a general summing up of this problem, we can state that the use of the ratio process in maintaining continuity in the event of a change in the weighting system is fairly good provided that the price-ratios do not fluctuate very violently, or the weight changes are not large.

B.—Ratio of aggregatives, using as basket of goods arbitrary year quantities.

This is covered in the general case of the weighted arithmetic mean, for we can reduce the present form of index number to the following form

$$I_{t0} = \frac{1}{\Sigma\, {}_iP_0\, {}_iQ_a} \Sigma \frac{{}_iP_t}{{}_iP_0} \cdot {}_iP_0\, {}_iQ_a.$$

The only qualification is that ${}_iv$ here involves the weight change comparing ${}_iP_0\, {}_iQ_b$ with ${}_iP_0\, {}_iQ_a$ in percentage forms.

5. *Simultaneous Changes—involving addition of some commodities and deletion of some others*

We have N commodities in the original index number, $i = 1, 2, \ldots, N$; for the new index number p of these commodities are to be deleted, and n other commodities are to be added. We use throughout this section the notations which have been introduced in the relevant foregoing sections.

A.—Arithmetic mean.

The same general results to be derived are applicable to all kinds of arithmetic mean index numbers, weighted or otherwise. Here only the results for the unweighted index numbers will be developed, and those for the weighted ones can be obtained as a simple extension.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The following relationship can be easily verified:

$$R = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = 1 + \frac{n\left(\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}\right) - p\left(\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}\right)}{N + n\frac{J_{t0}}{I_{t0}} - p\frac{K_{t0}}{I_{t0}}}$$

$$= 1 + E_c \text{ say.}$$

Now the error involved in the addition of commodities alone (E_a) is

$$\frac{n\left(\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}\right)}{N + n\left(\frac{J_{t0}}{I_{t0}}\right)},$$

and that involved in the deletion of commodities alone (E_d) is

$$-\frac{p\left(\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}\right)}{N - p\left(\frac{K_{t0}}{I_{t0}}\right)}.$$

The absolute magnitude of the first term in E_c,

$$\left| \frac{n\left(\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}\right)}{N + n\frac{J_{t0}}{I_{t0}} - p\frac{K_{t0}}{I_{t0}}} \right|,$$

is larger than $|E_a|$, while that of the second term in E_c,

$$\left| \frac{p\left(\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}\right)}{N + n\frac{J_{t0}}{I_{t0}} - p\frac{K_{t0}}{I_{t0}}} \right|,$$

is smaller than $|E_d|$.

If N is large compared to n and p, which is the situation we visualise in our discussion, then:

(i) If

$$\left| n\left(\frac{J_{T0}}{I_{T0}} - \frac{J_{t0}}{I_{t0}}\right) \right| > \left| \left(p\frac{K_{T0}}{I_{T0}} - \frac{K_{t0}}{I_{t0}}\right) \right|,$$

$$|E_c| > |E_a + E_d|.$$

(ii) In the reverse case,

$$|E_c| < |E_a + E_d|.$$

Generally we see that

$$|E_c| = |E_a\left[\frac{N + n\frac{J_{t0}}{I_{t0}}}{N + n\frac{J_{t0}}{I_{t0}} - p\frac{K_{t0}}{I_{t0}}}\right] + E_d\left[\frac{N - p\frac{K_{t0}}{I_{t0}}}{N + n\frac{J_{t0}}{I_{t0}} - p\frac{K_{t0}}{I_{t0}}}\right],$$

so that if N is large compared with nJ_{t0}/I_{t0} or pK_{t0}/I_{t0},

$$|E_c| \doteqdot |E_a + E_d.|$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Now we note that E_a and E_d are normally of opposite signs for more or less similar trends in the prices of the added and of the deleted commodities. It is thus possible to conclude that E_c is generally likely to be of smaller magnitude than either E_a or E_d.

B.—Geometric mean.

Similar results can be derived here as in the previous case, for we have

$$R = \left[\frac{\frac{J_{T0}}{I_{T0}}}{\frac{J_{t0}}{I_{t0}}}\right]^{\frac{n}{N+n-p}} \cdot \left[\frac{\frac{K_{t0}}{I_{t0}}}{\frac{K_{T0}}{I_{T0}}}\right]^{\frac{n}{N+n-p}}.$$

We recall that for addition of commodities, the ratio R, which we shall specify as R_a, is

$$R_a = \left[\frac{\frac{J_{T0}}{I_{T0}}}{\frac{J_{t0}}{I_{t0}}}\right]^{\frac{n}{N+n}}.$$

Also for deletion of commodities, the ratio R_d is

$$R_d = \left[\frac{\frac{K_{t0}}{I_{t0}}}{\frac{K_{T0}}{I_{T0}}}\right]^{\frac{p}{N-p}}.$$

The ratio in the case of addition and deletion of commodities, R_c, can be obtained in terms of R_a and R_d, thus

$$R_c = (R_a)^{\frac{N+n}{N+n-p}} \cdot (R_d)^{\frac{N-p}{N+n-p}}.$$

Now in this relationship we notice that

$$\frac{N+n}{N+n-p} > 1 \quad \text{and} \quad \frac{N-p}{N+n-p} < 1,$$

so that, since generally $R_a > 1$ and $R_d < 1$, or $R_a < 1$ and $R_d > 1$, provided that neither n nor p is large, R_c will be fairly small.

6. *Simultaneous Changes—involving a change in the base year and a change in the weighting system.*

This discussion is not concerned with the unweighted index numbers. The base-weighted ones have been dealt with in Section 1 under the change of the base year, for a change in the base year of the base-weighted index numbers automatically implies a change in the weighting system. Similarly current year weighted index numbers are not involved here, as by their nature such index numbers change their weighting system automatically year by year.

We are therefore left with the arbitrary year weighted index numbers.

A. Weighted mean of relatives.

We denote I'_{t0} and I'_{T0} as the index numbers which have been revised according to the new weighting system ${}_iP_b\,{}_iQ_b$ as compared with the old system ${}_iP_a\,{}_iQ_a$, and we study to what extent the ratio process can be used to derive I'_{tT} from I_{t0}/I_{T0}, I'_{tT} being the index number weighted by the new weighting system and based on year T instead of year 0. We can without loss of generality convert the weighting systems to percentage weighting, and denote

$${}_iW_a = \frac{{}_iP_a\,{}_iQ_a}{\Sigma\,{}_iP_a\,{}_iQ_a} \cdot 100$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and

$$ {}_iW_b = \frac{{}_iP_b\ {}_iQ_b}{\Sigma\ {}_iP_b\ {}_iQ_b}\ .\ 100. $$

We also use ${}_iv = {}_iW_b - {}_iW_a$.

(a) Arithmetic mean.

We can derive I'_{tT} from I_{tT} first from the following relationship:

$$ I'_{tT} = \Sigma\ \frac{{}_iP_t}{{}_iP_T}\ {}_iW_a + \Sigma\ \frac{{}_iP_t}{{}_iP_T}\ {}_iv $$

$$ = I_{tT} + \Sigma\ \frac{{}_iP_t}{{}_iP_T}\ {}_iv. $$

But

$$ 100\ \frac{I_{t0}}{I_{T0}} = I_{tT} - \Sigma\ \frac{{}_iP_t}{{}_iP_T}\left({}_iW_a - \frac{100\ {}_iW_a\ {}_id}{\Sigma\ {}_iW_a\ {}_id}\right) $$

where ${}_id = \frac{{}_iP_T}{{}_iP_0}$.

We therefore obtain

$$ 100\ \frac{I_{t0}}{I_{T0}} = I'_{tT} - \Sigma\ \frac{{}_iP_t}{{}_iP_T}\left({}_iW_a - \frac{100\ {}_iW_a\ {}_id}{\Sigma\ {}_iW_a\ {}_id}\right) - \Sigma\ \frac{{}_iP_t}{{}_iP_T}\ .\ {}_iv. $$

Now a comparison with the result derived in section 1 shows that the second term on the right-hand side of the above equation is the error involved in the departure from proportionality of the price-ratios ${}_id$. The third term is, of course, the error involved in the weight change.

We define the following:

$$ \frac{{}_iP_t}{{}_iP_T} = {}_i\rho; $$

$$ {}_iW_a - \frac{100\ {}_iW_a\ {}_id}{\Sigma\ {}_iW_a\ {}_id} = {}_i\delta. $$

It is clear that if there is strict proportionality between the prices in year T and those in year 0, ${}_i\delta$ would vanish for every commodity. Also $\Sigma\ {}_i\delta = 0$.

We finally obtain the following equation:

$$ 100\ \frac{I_{t0}}{I_{T0}} = I'_{tT} - N\ \sigma_\rho\ \sigma_\delta\ r_{\rho\delta} - N\ \sigma_\rho\ \sigma_v\ r_{\rho v}. $$

To assess the error involved, we consider the four situations that can arise:

(i) Increase in the price-trends, combined with weight changes in such a way that more weight will be given to those commodities whose prices have increased more, and less weight to those commodities which prices have increased less. In this situation obviously ρ will be positively correlated with δ, but negatively correlated with v, that is, $r_{\rho\delta}$ will be positive but $r_{\rho v}$ will be negative. If we denote the ratio $100I_{t0}/I_{T0}$ in the present case of weight-cum-base-year change as $R_{W;\,B}$, and the ratio $100\ I_{t0}/I_{T0}$ in the case of the base year change alone as R_B, then, although R_B is an under-estimate of I_{tT} involving base-change alone, $R_{W;\,B}$ if it is an under-estimate of I'_{tT}, will not under-estimate I'_{tT} as seriously as R_B does I_{tT}; on the other hand, $R_{W;\,B}$ may even over-estimate I'_{tT}, depending on the magnitudes of the two correlations.

For example, if we have $N = 10$, if ${}_i\rho$ ranges from 0·1 to 1·0, if the departure from proportionality is of the order of 20 per cent. either way, and if the weight change ranges from $-$ 15 per cent. to $+$ 15 per cent., then if no weight-change were introduced, but only the base year were shifted,

$$ 100\ \frac{I_{t0}}{I_{T0}} \doteqdot I_{tT} - 4\ r_{\rho\delta}. $$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

But if weight change were involved as well, then

$$100\,\frac{I_{t0}}{I_{T0}} \doteqdot I'_{tT} - 4\,r_{\rho\delta} - 3\,r_{\rho v}.$$

The maximum amount of under-estimate if $100\,I_{t0}/I_{T0}$ were used to estimate I'_{tT} would be obtained only if $r_{\rho\delta} = 1$ and $r_{\rho v} = 0$, and this would be the same as if there were no weight-change. But if $r_{\rho v}$ were not equal to zero, if it were, say, $-0{\cdot}4$, then the under-estimate would be only $-2{\cdot}8$ as compared with -4 in the event $r_{\rho v} = 0$.

There could be an over-estimate in this example if $|\,r_{\rho v}\,| > |\,r_{\rho\delta}\,|$. Say if $r_{\rho\delta} = 0{\cdot}2$ and $r_{\rho v} = -0{\cdot}8$, then there would be an over-estimate of $1{\cdot}6$. An over-estimate could of course also occur if $\sigma_v > \sigma_\delta$, i.e., if the weight changes were of larger variability than the departures from proportionality.

(ii) Increase in the price trends, combined with weight changes in favour of commodities which have increased less, and against commodities which have increased more. In this situation it is possible to see that $r_{\rho\delta}$ would still be positive, but that $r_{\rho v}$ would also be positive. The result is that $100\,I_{t0}/I_{T0}$ would yield a greater under-estimate of I'_{tT} than it does of I_{tT}. In other words, in this situation the ratio process is subject to a greater error than if no weight-changes were introduced.

(iii) Falling price trends, with weight changes favouring those commodities whose prices have fallen more and against those whose prices have fallen less. In this situation $r_{\rho\delta}$ is positive, and $r_{\rho v}$ is positive, and we obtain the same results as under (i).

(iv) Falling price trends, with weight changes giving more weights to those commodities which have fallen less in price and less weights to those which have fallen more in price. Here $r_{\rho\delta}$ is again positive but $r_{\rho v}$ is negative, yielding the same conclusions as those derived under situation (ii).

If prices fluctuate without any trend at all, then the above four situations are not really applicable. All that can be said is that the error incurred in using the ratio process for base-cum-weight change may be smaller than that incurred for base change alone; on the other hand it may be larger, the maximum error being an under-estimate of size

$$N\,\sigma_\rho(\sigma_\delta\,r_{\rho\delta} + \sigma_v\,r_{\rho v}).$$

(b) The geometric mean.

Similarly to the arithmetic mean, we can obtain this relationship for the geometric mean index number:

$$100\,\frac{I_{t0}}{I_{T0}} = I'_{tT} \cdot \sqrt[100]{\Pi\left(\frac{{}_iP_T}{{}_iP_t}\right)^{i^v}}.$$

A comparison with the case of weight change alone shows that this is identical with the relationship derived there. For since

$$R = \frac{\dfrac{I'_{T0}}{I_{T0}}}{\dfrac{I'_{t0}}{I_{t0}}} = \sqrt[100]{\Pi\left(\frac{{}_iP_T}{{}_iP_t}\right)^{i^v}},$$

$$I'_{T0} \cdot \frac{I_{t0}}{I_{T0}} = I'_{t0} \cdot \sqrt[100]{\Pi\left(\frac{{}_iP_T}{{}_iP_t}\right)^{i^v}}.$$

This is the relation applicable to weight change alone, and if base change to year T is introduced, then I'_{T0} becomes 100, and I'_{t0} becomes I'_{tT}.

Therefore the case of simultaneous change of base year and weighting system is precisely that of the change due to weighting system by itself. This latter has already been discussed.

(c) Ratio of aggregatives.

Again the same discussion and the same conclusions apply here as for the arithmetic mean so long as we remember that the ${}_iv$ corresponding to that case is now the difference between ${}_iP_T\,{}_iQ_b$ and ${}_iP_T\,{}_iQ_a$ in percentage terms.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

7. *Simultaneous Change—involving a change in the base year and addition of some commodities and deletion of some others*

Here I_{t0} and I_{T0} are the old index numbers before the introduction of the base and the commodity changes, and I'_{tT} is the new index number with the changes introduced. We study to what extent the ratio $100\, I_{t0}/I_{T0}$ can be used to estimate I'_{tT}.

A.—Arithmetic mean.

Again we shall study only the case of the unweighted arithmetic mean. The results for the weighted arithmetic mean index numbers are similar to those of the unweighted one.

A reference to the section dealing with the commodity change (that is, with the addition and deletion of commodities) gives us the following:

$$100\,\frac{I_{t0}}{I_{T0}} = 100\,\frac{I'_{t0}}{I'_{T0}}\,(1 - E_{a;\,d})$$

where I'_{t0} and I'_{T0} are the index numbers involving the commodity change only, and $E_{a;\,d}$ is the error incurred.

Also as an analogy with the case of base year change alone, we can derive $100\, I'_{t0}/I'_{T0}$ thus:

$$100\,\frac{I'_{t0}}{I'_{T0}} = I'_{tT} - \frac{100}{N + n - p}\,\Sigma\,\frac{{}_{i'}P_t}{{}_{i'}P_T}\left(1 - \frac{{}_{i'}a}{\bar{a}}\right)$$

where i' covers the range of commodities taking into consideration the added and the deleted ones. If we use E_b to denote the error incurred in the event of base change, we obtain

$$100\,\frac{I'_{t0}}{I'_{T0}} = I'_{tT}(1 - E_b).$$

Hence finally

$$100\,\frac{I_{t0}}{I_{T0}} = I'_{tT}(1 - E_b)\,(1 - E_{a,\,d}).$$

In studying the error here, it is to be noted first of all that E_b implies an under-estimate, i.e., E_b is positive. We therefore see that if $E_{a;\,d}$ is negative, then the error involved in the present case ($E_{b;\,c}$ say), is smaller than either E_b or $E_{a;\,d}$ if both are of about the same order of magnitude. But if either one of these is very small and the other one fairly large, then $E_{b;\,c}$ is likely to be larger than the former, but it will be smaller than the latter.

B.—Geometric mean.

The problem of the geometric mean is relatively simpler. For the unweighted geometric mean and for the weighted geometric mean weighted by arbitrary year values it is easy to verify that the error involved in the base-cum-commodity change is identical with that involved in commodity change alone. For the weighted geometric mean index numbers weighted by base-year values and by current-year values, the same results can be obtained as in the general case of the arithmetic mean.

C.—Ratio of aggregatives.

Here, too, arbitrary year weighting can be seen to be identical in the present situation and in the situation of commodity-change alone, while for the base weighted and current weighted ratios, results can be deduced similar to those of the arithmetic mean.

8. *Simultaneous Change, involving change in the weighting system and in the commodity-coverage*

Since the current weighted index numbers are not applicable here, and since for the base weighted index numbers to change the weighting system would involve also a change in the base year, we are in this section again faced only with the problem of the arbitrary year weighted index numbers.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In the present situation the changed weights of the deleted commodities k would be zero throughout; also the weights of the added commodities j in the new index number would all be positive.

We shall find it convenient to use percentage weighting systems, and we shall follow the same symbols as used in section (6) with regard to weight changes.

A.—Weighted arithmetic mean.

Following on what has been said above, we see that

$$I'_{T0} = \Sigma \frac{{}_iP_T}{{}_iP_0} \cdot {}_iW_b + \Sigma \frac{{}_jP_T}{{}_jP_0} \cdot {}_jW_b - \Sigma \frac{{}_kP_T}{{}_kP_0} \cdot {}_kW_b$$

$$= \Sigma \frac{{}_iP_T}{{}_iP_0} \cdot ({}_iW_a + {}_iv) + \Sigma \frac{{}_jP_T}{{}_jP_0} \cdot {}_jW_b.$$

The i original commodities consist of two portions: k of these i are involved in the deletion, leaving us with $h = i - k$ commodities which are unaffected by either addition or deletion. Now ${}_kv = -{}_kW_a$, and remembering that $I_{T0} = \Sigma ({}_iP_T/{}_iP_0)\,{}_iW_a$, we obtain

$$\frac{I'_{T0}}{I_{T0}} = 1 + \Sigma \left(\frac{{}_hP_T}{{}_hP_0} \Big/ I_{T0}\right) {}_hv + \Sigma \left(\frac{{}_jP_T}{{}_jP_0} \Big/ I_{T0}\right) {}_jW_b - \Sigma \left(\frac{{}_kP_T}{{}_kP_0} \Big/ I_{T0}\right) {}_iW_a.$$

Hence

$$R = \frac{\dfrac{I'_{T0}}{I_{T0}}}{\dfrac{I'_{t0}}{I_{t0}}}$$

$$= 1 + \frac{\Sigma {}_hv \left(\dfrac{{}_hP_T}{{}_hP_0} \Big/ I_{T0} - \dfrac{{}_hP_t}{{}_hP_0} \Big/ I_{t0}\right) + \Sigma {}_jW_b \left(\dfrac{{}_jP_T}{{}_jP_0} \Big/ I_{T0} - \dfrac{{}_jP_t}{{}_jP_0} \Big/ I_{t0}\right) - \Sigma {}_kW_a \left(\dfrac{{}_kP_T}{{}_kP_0} \Big/ I_{T0} - \dfrac{{}_kP_t}{{}_kP_0} \Big/ I_{t0}\right)}{1 + \Sigma {}_hv \left(\dfrac{{}_hP_t}{{}_hP_0} \Big/ I_{t0}\right) + \Sigma {}_jW_b \left(\dfrac{{}_jP_t}{{}_jP_0} \Big/ I_{t0}\right) + \Sigma {}_kW_a \left(\dfrac{{}_kP_t}{{}_kP_0} \Big/ I_{t0}\right)}$$

$$= 1 + E_{w;\,c} \text{ (say).}$$

We shall here consider two situations, as follows:

(i) Where the weights of the added commodities are mainly contributions from those of the deleted commodities, and the weights of the commodities h are re-allocated among themselves. In this situation,

$$\Sigma {}_hv \left(\frac{{}_hP_t}{{}_hP_0} \Big/ I_{t0}\right)$$

will be small, and in view of the fact that both

$$\Sigma {}_jW_b \left(\frac{{}_jP_t}{{}_jP_0} \Big/ I_{t0}\right)$$

and

$$\Sigma {}_jW_a \left(\frac{{}_kP_t}{{}_kP_0} \Big/ I_{t0}\right)$$

are necessarily positive, we can, by reference to the results of sections 2, 3 and 4, derive the following relationship:

$$|\, E_{w;\,c} \,| < |\, E_w + E_a + E_d \,|$$

where E_w is the error due to weight change alone of the h commodities, E_a the error due to the addition of the j commodities, and E_d the error due to the deletion of the k commodities.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Now generally it can be seen from the results of sections 2 and 3 that E_a and E_d will be of opposite signs for more or less similar trends in the prices of the added and of the deleted commodities. These two errors will actually tend to cancel each other out, so that $E_{w;\,c}$ will be smaller than E_w or in any case not very different from it.

(ii) The second situation is that where the weights of the deleted commodities may not be sufficient to make up for the weights of the added commodities, so that the weights of the h commodities (or at any rate of some of them) have to be also reduced for the purpose. Here

$$\Sigma\, {}_hv \left(\frac{{}_hP_t}{{}_hP_0} \Big/ I_{t0} \right)$$

is likely to be negative.

There is, however, a limit to the magnitude of this quantity. For one of the underlying principles in our study is that the number of added and of deleted commodities should not be large, and also that they should not assume such unduly large proportion, and therefore the basic commodity coverage should not be so changed, that the index number loses its original identity. If it does, then there does not appear to be any purpose in maintaining the continuity of two essentially dissimilar series of index numbers.

In any case the inequality derived above would be inapplicable only if

$$\Sigma\, {}_hv \left(\frac{{}_hP_t}{{}_hP_0} \Big/ I_{t0} \right) + \Sigma\, {}_jW_b \left(\frac{{}_jP_t}{{}_jP_0} \Big/ I_{t0} \right) < \sigma$$

or

$$\Sigma\, {}_hv \left(\frac{{}_hP_t}{{}_hP_0} \Big/ I_{t0} \right) + \Sigma\, {}_jW_a \left(\frac{{}_kP_t}{{}_kP_0} \Big/ I_{t0} \right) < 0$$

or both. These conditions are expressible thus:

$$\Sigma\, {}_hW_b \cdot \frac{{}_hP_t}{{}_hP_0} + \Sigma\, {}_jW_b \cdot \frac{{}_jP_t}{{}_jP_0} < \Sigma\, {}_hW_a \cdot \frac{{}_hP_t}{{}_hP_0}$$

and

$$\Sigma\, {}_hW_b \cdot \frac{{}_hP_t}{{}_hP_0} < \Sigma\, {}_hW_a \cdot \frac{{}_hP_t}{{}_hP_0} - \Sigma\, {}_kW_a \cdot \frac{{}_kP_t}{{}_kP_0}$$

The first of these conditions implies that the weights of the added commodities are at least made up for as a result of the weight re-allocation among the h commodities. The second condition implies that all the weights of the deleted commodities and more besides are used towards the weight re-allocation of the h commodities.

These implications are at variance with the basis of our argument, and if they do arise, they strictly fall outside the scope of our study.

B.—Geometric mean.

With this index number it can be seen that

$$I'_{T0} = 100. \sqrt[100]{\frac{\Pi \left(\frac{{}_iP_T}{{}_iP_0} \right)^{{}_iW_a} \Pi \left(\frac{{}_hP_T}{{}_hP_0} \right)^{{}_hv} \Pi \left(\frac{{}_jP_T}{{}_jP_0} \right)^{{}_jW_b}}{\Pi \left(\frac{{}_kP_T}{{}_kP_0} \right)^{{}_kW_a}}}.$$

From this we can derive the following relationship:

$$R_{w;\,c} = \frac{\frac{I'_{T0}}{I_{T0}}}{\frac{I'_{t0}}{I_{t0}}} = \sqrt[100]{\frac{\Pi \left(\frac{{}_hP_T}{{}_hP_t} \right)^{{}_hv} \cdot \Pi \left(\frac{{}_jP_T}{{}_jP_t} \right)^{{}_jW_b}}{\Pi \left(\frac{{}_kP_T}{{}_kP_t} \right)^{{}_kW_a}}}.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A reference to sections 2, 3 and 4 yields:

$$R_{w;\,c} = R_w \cdot R_a \cdot R_d \left(\frac{I_{T0}}{I_{t0}}\right)^{\frac{\Sigma_j W_b - \Sigma_k W_a}{100}},$$

where R_w, R_a and R_d are the ratios for weight change, for addition of commodities and for deletion of commodities separately.

It can be seen that, as before, if the weights of the added commodities do not differ very much from those of the deleted commodities, then

$$R_{w;\,c} \doteqdot R_w \cdot R_a \cdot R_d.$$

Also from sections 2 and 3 it is clear that generally R_a and R_d differ from 1 in opposite directions for more or less similar price trends of the added and of the deleted commodities, so that $R_a \times R_d$ will be nearly 1. That leaves us with almost the same conclusion as for the arithmetic mean, namely, that $R_{w;\,c}$ will be approximately the same as R_w.

C.—Ratio of aggregatives.

Once again similar conclusions can be derived here as for the arithmetic mean. It is only pertinent to mention that the percentage weights W refer to the weighting system $P_0\,Q_b$ or $P_0\,Q_a$ whichever is applicable.

9. *Simultaneous Change, involving a shift in the base year, a change in the weighting system, and a change in the commodity coverage.*

This, the last case to be studied, of course involves only weighted index numbers.

A.—Arithmetic mean and ratio-of-aggregatives index numbers.

These two major types of index numbers can, from what we have seen before, be discussed together, for the base weighted arithmetic mean indices are identical to the Lespeyre ones, and the arbitrary and current year weighted indices of the aggregative-ratio type are similar to those of the arithmetic mean type.

Also the base weighted and the arbitrary year weighted index numbers can be studied together, for as far as the weighting systems are concerned, we can regard the old base year 0 (with percentage weights W_0) as the old arbitrary year (a), and the new base year T (with percentage weights W_T) as the new arbitrary year (b).

(a) Base-weighted and arbitrary-year weighted index numbers. The solution of the problem here is similar to that studied under Section 7, dealing with a simultaneous change in base year year and a change in the commodity coverage. We derive, in fact, the following relationship:

$$100\,\frac{I_{t0}}{I_{T0}} = I'_{tT}(1 - E_b)\,(1 - E_{w;\,c}),$$

where E_b is the error involved in the base year change, and $E_{w;\,c}$ is, from Section 8, that involved in the weight-cum-coverage change.

Now while $E_{w;\,c}$ is similar in both the base year and arbitrary year weighted index numbers, E_b is not necessarily similar. For in the arbitrary year weighted index numbers we see from section 1 that the error is dependent on the deviation from strict proportionality of the price ratios P_T/P_0, and therefore, in the same way as the unweighted arithmetic mean index number, E_b is likely in general to be an under-estimate. On the other hand, for base weighted index numbers, E_b is dependent on the departure from strict proportionality of the quantity ratios Q_0/Q_T. Thus if price trend has been generally upwards, and if the consumption has also been following an upward trend, then E_b would be an under-estimate, but if consumption has been downward generally, E_b would be an over-estimate.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B.—Geometric mean.

It would be preferable here to deal with the base weighted and the arbitrary year weighted index numbers separately, because the error involved in the base change is very different in the two cases.

(*a*) Using arbitrary year weights. This is dealt with first because it is the simpler of the two. As in section 7, for this index number the error involved in the situation of base-cum-weight-cum-commodity change is identical with that involved in the situation of weight-cum-commodity change only. For we can easily obtain:

$$I'_{tT} = I_{tT} \cdot \sqrt[100]{\frac{\Pi\left(\frac{{}_hP_t}{{}_hP_T}\right)^{{}_hv} \cdot \Pi\left(\frac{{}_jP_t}{{}_jP_T}\right)^{{}_jW_b}}{\Pi\left(\frac{{}_kP_t}{{}_kP_T}\right)^{{}_kW_a}}}$$

$$= \frac{I_{tT}}{R_{w;\,c}}.$$

Here I' refers to the index number with the weight and commodity changes introduced, and I to the index number based on the original weights and commodity coverage. $R_{w;\,c}$ is identical with that obtained under Section 8 (B).

Since for this type of index number the following obtains:

$$100\,\frac{I_{t0}}{I_{T0}} = I_{tT},$$

we finally derive

$$100\,\frac{I_{t0}}{I_{T0}} = I'_{tT} \cdot R_{w;\,c},$$

which is what we require.

(*b*) Using base year weights. Here the situation is somewhat complicated by the fact that the base change introduces an error as well. In fact it can be seen that

$$100\,\frac{I_{t0}}{I_{T0}} = I_{tT} \cdot R_b$$

where R_b is the ratio derived under the condition of base change. In a similar manner to (*a*) above, we derive this relationship:

$$100\,\frac{I_{t0}}{I_{T0}} = I'_{tT}\left(\frac{R_{w;\,c}}{R_b}\right).$$

References

SENG, YOU POH (1956) "Some theory of index numbers", *J.R. Statist. Soc. A*, **119**, 312–332.

TIPPETT, L. H. C. (1925) "On the extreme individuals and the range of samples taken from normal population", *Biometrika*, **27**, 364–387.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Note on a Visit to Moscow for the Third Soviet Mathematical Congress

By M. S. Bartlett

An invitation from the Academy of Sciences of the U.S.S.R. to attend the Third Soviet Mathematical Congress, held in Moscow from June 25th to July 4th, 1956, recently gave me an opportunity to meet Russian probabilists and mathematical statisticians, and to learn more about the present status and scope of probability theory and mathematical statistics within the U.S.S.R. This note has been prepared in response to a suggestion made to me that a brief report of my visit would also be of interest to others.

The papers given at the Congress were divided into thirteen sections:

1. The theory of numbers.
2. Algebra.
3. Differential and integral calculus.
4. The theory of functions.
5. Functional analysis.
6. The theory of probability.
7. Topology.
8. Geometry.
9. Mathematical logic and foundations of mathematics.
10. Numerical mathematics.
11. Mathematical problems of mechanics.
12. Mathematical problems of physics.
13. History of mathematics.

These topics are rather similar to those listed in the last International Congress of Mathematicians at Amsterdam, although the omission of Statistics from the title of Section 6 made it difficult to know in advance whether mathematical statistics, as distinct from the mathematical theory of probability, was to be included. The total number of participants within the Soviet Union was large, over 1000 in all, though this would include many university students. The number of foreign visitors was about 60, of whom I should say more than half were delegates from the East European countries and China. The rest of us were individual guests, including two Americans, C. R. Rao from India, and Miss Cartwright and myself from England. I gathered that H. Cramér, of Sweden, and A. Rényi, of Hungary, were among those invited but unable to come.

There was of course considerable overlap of the times of papers in the different Sections, so that attendance had largely to be confined to one's own Section, although papers in Sections 10–13 were also sometimes of interest to me. In Section 6 the titles of the individual papers indicated that applications of probability theory and topics in mathematical statistics were in fact being included. These papers included a few on specific physical applications and one on quality control. Perhaps the most coordinated group of papers was that on information theory (in Shannon's sense, as used in the theory of communication) and the first in the group, by A. N. Kolmogorov, I. M. Guelfand and A. M. Yaglom, entitled "The quantity of information and the entropy for continuous distributions with a finite or infinite number of dimensions", was presented by Kolmogorov to a large audience. It was largely concerned, as were some of the other papers in this group, with the rigorization and extension of some of the basic definitions and formulae. Kolmogorov played a prominent part in the Congress, presenting papers in at least three different Sections and taking a leading part in many of the discussions.

In the theory of stochastic processes a fundamental paper was given by E. B. Dynkin on "New methods in the theory of Markov processes". It was difficult, even with the aid of an interpreter, for me to follow papers given in Russian, but Dynkin was apparently attempting to classify all Markov processes in continuous time by means of their infinitesimal operators. He made use not

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

only of semi-group theory but also of the probabilistic notion of strong Markov processes, defined, roughly, as stochastic processes for which the Markov property held at random times (defined without reference to the future) as well as at fixed times. He claimed in particular to have obtained Feller's results for one-dimensional diffusion processes under less restrictions than Feller, and to be studying multi-dimensional diffusion processes by these new methods. (This work is included in that to be published in a new Russian journal on probability, of which further details are given in the Appendix.)

An interesting paper by B. V. Gnedenko, N. V. Smirnov and I. I. Gikhmann on non-parametric methods in statistics was presented by Gnedenko, who mentioned to me privately that he was publishing a book shortly (with Gikhmann) on this topic.

C. R. Rao gave two papers, one on estimation theory and the definition of consistency, and the other on discriminant function theory for growth curves. It might be noted that the latter paper by Rao, and my own paper on applications of probability and statistical theory in biology, were the only papers in Section 6 with a biometrical flavour. The primarily mathematical nature of the Congress should be borne in mind before assigning any significance to this, but nevertheless it appeared to me that the main interests of the Russians lay in pure theory and physical applications. This is probably largely a matter of inclination and custom, and Russian contributions to the fundamentals of probability theory do not need to be emphasized here, though it is possible that any repercussions within the U.S.S.R. on the use of statistics in biology arising from earlier biological controversies had not yet entirely subsided. However, I found the whole academic atmosphere at the Congress quite free and unrestricted; and in particular my own survey paper was translated into Russian and read paragraph by paragraph with my English version. This is consistent with the recent recognition in the U.S.S.R. of mathematical statistics as an important subject in its own right, and of the value of applications of probability theory to many branches of science and technology.

The announcement of a new journal entitled *The theory of probability and its applications*, open to contributions from outside the Soviet Union, is a further indication of an unrestricted atmosphere. Although the word Statistics does not explicitly occur in the title of this journal, it will be seen from the official details reproduced in the Appendix below that topics in mathematical statistics are to be included. The first number will contain papers by Khintchine, Cramér, Dynkin, Rényi, Dobrušin, Takacs, Tumanyan, Mihoc and Vladimirov.

Further private discussions with the Russians were extremely cordial. I obtained from Kolmogorov full particulars of the mathematics courses and seminars held at Moscow. Among physical applications of interest to the Moscow mathematicians, who included Yaglom and Guelfand, were the theory of turbulence (though it did not appear that Kolmogorov himself was working any longer in this field) and the use of probability in quantum mechanics. Gnedenko, who is at Kiev, was interested in applications of statistical techniques in medicine, especially in the study of heart disease. Some of the younger members, Prokhorov, for example, were very helpful during some of these discussions, as they could speak English.

I also took the opportunity of asking Kolmogorov about some of the well-known Russian probabilists. Both Slutsky and Romanovsky were now dead. Bernstein, who has retired, and Khintchine, who I heard was ill, were not at the Congress. Smirnov, though apparently retired, was quite active at the Congress, and presented me with a book, of which he was one author, on mathematical statistics in relation to technical applications. I understood from Gnedenko that the only previous book on mathematical statistics was the one by Romanovsky; Bojarsky had written books on demographic and medical statistics.

I did not attempt to pursue the relation of mathematical statistics with statistical experimentation in agriculture and other biological fields (although I had referred to English contributions in these general fields in my paper) as it did not appear likely that those present during these discussions had any particularly close contact with such research. However, in view of my own interest in stochastic models in animal ecology (a topic to which I had also briefly referred in my paper), I was especially pleased during my stay in Moscow to be able to meet G. F. Gause, who has done such important work in the ecological field. He is now Deputy-Director of the Institute of Antibiotics, Academy of Medical Sciences. I had a very stimulating conversation with Gause, who spoke excellent English, and seemed especially interested in modern statistical work within his general province e.g. in bio-assay.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The various official excursions and visits arranged during the Congress included a visit to the high-speed electronic computer housed in the Institute of Exact Mechanics and Computing Technique. It was interesting to find in Moscow both this installation and (for example, in many shops) the simple abacus, two extremes in the evolutionary range of computing machines.

APPENDIX

The Academy of Sciences of the U.S.S.R. is undertaking in 1956 the publication of a journal "Theoriya veroyatnostei i ee primeneniya" ("The Theory of Probability and its applications").

Editorial board: Bernstein, S. N., Gnedenko, B. V., Zolotarev, V. M., Kolmogorov, A. N. (Editor in Chief), Korolyuk, V. S. (Assistant Editor), Linnik, Yu. V., Sarymsakóv, T. A., Smirnov, N. V., Khintchine, A.

The journal is to contain original articles and short reports on the theory of probability, general problems of mathematical statistics as well as the application of the theory of probability to natural sciences and technology.

Articles of the latter kind will be accepted only if the mathematical methods applied are essentially new.

Articles and short reports will be published either in Russian with summaries in English, German or French, or else in English, German or French with summaries in Russian.

The authors will be sent 50 copies of their articles free of charge.

The journal will also include reviews of books on the Theory of Probability and Mathematical Statistics as well as periodic lists of publications received.

Authors are requested to address their contributions: Moscow, W-234, MGU—Department of Mathematics and Mechanics, or Kiev, Kalinin Square 6, Institute of Mathematics of the Academy of Sciences of the Ukrainian SSR.

The journal appears quarterly. The price for 4 issues is 45 roubles.

Addresses of the agencies where subscriptions should be sent to (including back numbers of the journal) are supplied. For example:

England—Collet's Export-Import Department, 44–45, Museum Street, London, W.C.1.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REPORT OF THE COUNCIL

For the FINANCIAL YEAR *ended December 31st, 1955, and for the* SESSIONAL YEAR *ended June 20th, 1956, presented at the* ONE HUNDRED AND TWENTY-SECOND ANNUAL GENERAL MEETING *of the* ROYAL STATISTICAL SOCIETY, *held at the London School of Hygiene and Tropical Medicine, W.C.1, on June 20th, 1956.*

Number of Fellows

A FURTHER increase in the number of Fellows took place during the year, raising the total to over 2,600. The actual number at December 31st, 1955, was 2,619 that figure representing a net gain of 80. During 1955 182 Fellows were elected or restored to the roll and 102 were lost by death, resignation or default.

Further elections in the early part of 1956 brought the total number of Fellows at the end of April to 2,670.

At the end of 1955, Honorary Fellows numbered 31, of whom 7 were Presidents of other Societies. New elections raised the previous total by 6, but 3 Honorary Fellows, all of long standing, died since the last Annual Report.

Representatives of corporate bodies or institutions are included in the total of 2,619 ordinary Fellows. Their number was 117.

Number of Fellows (excluding Honorary Fellows)

Calendar Year	*Losses*		*Elected or Restored to the Roll*	*On the Roll at December 31st*
	By Death	*By Resignation or Default*		
1944	13	16	159	1,269
1945	18	11	177	1,417
1946	24	32	287	1,648
1947	14	61	263	1,836
1948	17	79	238	1,978
1949	16	87	189	2,064
1950	10	83	222	2,193
1951	21	93	228	2,307
1952	15	89	221	2,424
1953	10	91	165	2,488
1954	14	106	171	2,539
1955	19	83	182	2,619

Losses by Death

The Council records with regret the loss by death during the year ended April 30th, 1956, of the undermentioned Fellows:

	Date of Election
Borel, Professor Emile (Honorary Fellow)	1934
Methorst, Professor H. W. (Honorary Fellow)	1920
*p*Shirras, Professor G. Findlay (Honorary Fellow—ordinary Fellow 1911–1920)	1920
Barou, Noah, Ph.D.(Econ.)	1943
Bateman, Gertrude Isabella, B.A., Ph.D.	1949
*Chamberlain, John Alfred	1927
Dalberg, Rudolph, Dr.phil.(Econ.), Dr.Jur.	1946
Denning, Sidney Edward, F.C.I.S.	1923
Furlan, Professor Louis V.	1922
Gardner, Benjamin	1945
Gaughan, Louis	1943
Goodall, George, M.A.	1938
*Guy, John Henry	1948
*c**Jevons, Herbert Stanley, M.A., B.Sc.	1907
c Lyle, Philip	1946
Nicholls, Helen	1945
Orwin, Charles Stewart, M.A., D.Litt.	1917
*Smith, Stanley George	1905
Smith, Professor Wilfred, M.A.	1944
Vibart, Hugh Henry Rose, M.A., B.Litt.	1923
*Watson-Hyatt, Ronald William, O.B.E., E.D., B.A., F.C.I.S.	1946

* Life Fellow. *c* Served on Council. *p* Contributed to *Proceedings*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Professor Emile Borel was a mathematician of great distinction, well known for his interpretations and criticisms of Einstein's theories as well as for his own work on the theory of functions and the calculus of probabilities. He held many high honours, political as well as scientific, in France, and his strong international interests found expression in serving as President of the World League of Nations Union and as a Vice-President of both the European Economic and Customs Union and the International Council of Scientific Unions. He was an Honorary Member of the International Institute of Statistics and was elected an Honorary Fellow of the Society in 1934.

Professor H. W. Methorst, who was 87 when he died, entered the Dutch official statistical service in 1895 and rose with its development until in 1927 he was Director-General of the Central Bureau of Statistics, an office which he held until his retirement in 1937. As Secretary General of the International Institute of Statistics he participated in the arrangements for and conduct of the Jubilee meeting of that body at the Royal Statistical Society's Centenary in 1934. His election as an Honorary Fellow in 1920 was a recognition of his prominence in Dutch statistics, where his particular interests lay chiefly in problems of population and health.

Professor G. Findlay Shirras was widely known in British and international statistical circles although his active life from 1909 was mainly spent in India. Appointed in that year Professor of Economics at Dacca College, he was attached to the Finance Department of the Government of India from 1910 to 1913, and in 1914 was given the newly-created post of Director of Statistics, which he held till 1921. In this capacity he attended the British Empire Statistical Conference in London in 1919–20, and in virtue of this position he was elected an Honorary Fellow of the Society in 1920. Outstanding among his services to India was the Statistical Memorandum which secured for that country her recognition as one of the "eight states of chief industrial importance" and so entitled to one of the non-elective seats on the Governing Body of the International Labour Office.

By the death of Mr. Philip Lyle in July, 1955, the Council lost a devoted member and the Society a stalwart supporter whose services were of the greatest value and variety. Though he became a Fellow only in 1946, his interest in statistics stretched back through many years. Elected to the Council in 1947, he was on the Executive Committee from 1948–1950 and since 1950 had been a Joint Honorary Secretary. He was a Vice-President for the years 1952–53 and 1953–54, a member of the Industrial Applications Section Committee from 1950 to 1955 and of the Committee for *Applied Statistics* from its inception in 1951 until 1955. He took a leading part in the re-housing of the Society and was Chairman of the House Committee during its existence, 1953–54. All who came into contact with him over the Society's affairs held him in affection as well as in respect. His unsparing services to the Industrial Applications Section and his part in the launching of *Applied Statistics* will perhaps be most widely remembered, but rarely can the Society have had to mourn the loss of one so humbly ready to serve in whatever channel offered, so disinterested and so wise in consultation. There is no branch of the Society's activities in which that loss is not keenly felt.

Professor Stanley Jevons furthered the cause of statistics in many countries, especially in India. A member of Council in the year 1932–33 and again in 1936–37, he also took part in discussions on papers connected with the causes of industrial fluctuations, reading a paper on that subject himself in 1933 and sharing discussions right up to 1953. His published works had a wide range—cotton, coal, finance, tenancy law and estate management, petrology, mineralogy and political economy. His idealism was as evident as his scientific knowledge, and he was well known for his championship of Ethiopia after Mussolini's attack on that country. It was characteristic of the breadth of his interest that his last published work should be entitled *Economic Equality in the Co-operative Commonwealth* .

Dr. C. S. Orwin, first Director of the Agricultural Economic Research Institute in the University of Oxford, had a long record of service to agriculture, having begun by studying at Wye College in its early years. After short periods with an estate agent and in business, he returned to Wye as lecturer till 1906, when he was invited to manage and develop estates in Lincolnshire. There, for the next seven years, he gained the experience which underlay the great work he later carried out so successfully, and which also determined the practical basis of his approach to every problem. During this period he developed methods of book-keeping and cost-accounting suited to British farming complexities, yet simple enough to be grasped by the ordinary farmer. At Oxford, from

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

1913, he had to set up and guide the development of a new type of research institution in a new subject. By the end of his thirty-three years as Director, that Institute was an integral part of the University and a recognized national centre for work in agricultural economics; all this was the achievement of his personality, patience and persistence. His own lucid, succinct style of writing was a model for all statisticians and, while his direct contributions to knowledge were substantial, his indirect influence through the hundreds of post-graduate students trained at the Institute was immeasurable.

Sir Ronald Fisher, F.R.S.

The Council wishes to place on record the satisfaction it felt on the award to Sir Ronald Fisher of the Copley Medal of the Royal Society. There is no need to emphasize the high honour conveyed by this distinction to Sir Ronald and the Council is glad to express its pleasure on this further acknowledgment of his work.

Honorary Fellowship

At the Ordinary General meeting on June 22nd, 1955, six statisticians of international reputation were elected to Honorary Fellowship, their names being as follows:—

ProfessorMaurice Fréchet	Dr. Lowell J. Reed
Dr. Gunnar Jahn	Professor J. F. Steffensen
Professor Paul Lévy	Dr. Edwin B. Wilson

During the past Session the Council has again considered the desirability of extending the Society's roll of Honorary Fellowship and has decided to recommend four more well-known statisticians for election.

It is with regret that the Council records the loss by death of three distinguished statisticians who had long been Honorary Fellows of the Society, and to whom reference is made in the preceding section of the Report.

Vice-Presidents

For the Session 1955–56 the President appointed as Vice-Presidents of the Society Professor D. G. Champernowne, Dr. D. J. Finney, Professor M. G. Kendall and Sir George Maddex.

Meetings of the Society

As in previous years, the Council is glad to acknowledge its gratitude to the Board of Management and the Dean of the London School of Hygiene and Tropical Medicine for the use of their lecture theatre for the Ordinary General Meetings of the Society and the meetings of the Research Section. The Council is, this year, especially indebted for this privilege as the Authorities of the School decided to include the Society amongst those learned bodies to whom the use of the lecture theatre is granted without charge.

During the Session the following papers were read at the Ordinary General Meetings:

1955	
November 16th	D. J. Finney. The Statistician and the Planning of Field Experiments.
December 21st	H. Wold. Causal Inference from Observational Data.
1956	
January 18th	E. S. Pearson. Presidential Address. Some Aspects of the Geometry of Statistics. The Use of Visual Presentation in Understanding the Theory and Application of Mathematical Statistics.
February 15th	P. E. Hart and S. J. Prais. The Analysis of Business Concentration: A Statistical Approach.
March 21st	R. H. S. Phillips. United Kingdom Indices of Wholesale Prices, 1949–55.
April 18th	B. P. Emmett The Television Audience in the United Kingdom.
May 16th	Dorothy Cole and J. E. G. Utting. Estimating Expenditure, Saving and Income from Household Budgets.

The number of Fellows and visitors present has varied between 60 and 219.

In recent years the Council has referred in its Annual Report to the difficulty of securing offers of suitable papers for reading. Possibly as the result of these references the difficulty in the past Session has been that the papers submitted or offered have exceeded in number the opportunities

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

for presenting them. This is not a cause for regret; on the contrary, the Council welcomes the experience.

The Council again reminds authors that papers of undue length are strongly discouraged. Long papers do not necessarily ensure the success of the Society's meetings and are expensive to reproduce in the *Journal*. With the rise in printing costs which took place at the beginning of 1956 it is even more necessary to emphasize the merit of brevity, and in circumstances where the Council has to make a selection from the papers offered, their length must be a consideration.

The Research Section

Dr. H. E. Daniels was elected Chairman of the Research Section for the Session, and Dr. C. A. B. Smith was re-elected Honorary Secretary. Other members of the Committee were: Mr. N. T. J. Bailey, Professor G. A. Barnard, Dr. G. E. P. Box, Mr. J. Durbin, Mr. J. M. Hammersley, Dr. N. L. Johnson, Dr. S. Vajda, and Mr. A. M. Walker, with Dr. P. Armitage and Professor M. S. Bartlett appointed by the Council and Dr. J. O. Irwin an *ex officio* member as Editor of the *Journal*, Series B.

The papers read during the Session have been as follows:

1955 November 2nd	F. N. David and N. L. Johnson. Some Tests of Significance with Ordered Variables.
1956 January 4th	P. M. Grundy, M. J. R. Healy and D. H. Rees. Economic Choice of the Amount of Experimentation.
March 14th	C. L. Mallows. Generalizations of Tchebycheff's Inequalities.
May 2nd	Violet Cane. Some Statistical Problems in Experimental Psychology.

The Industrial Applications Section

The Section Committee which co-ordinates the functions of the separately organized Local Groups and links them with the Society has been constituted as follows: (*a*) appointed by Council: Mr. D. J. Desmond, Mr. G. W. Sears, Mr. N. J. Squirrell, Mr. A. W. Swan, Mr. E. D. van Rest, Mr. P. J. Stanley, and the late Mr. Philip Lyle; (*b*) nominated by the 1954–55 Section Committee: Miss J. Keen, Mr. M. Whyte, Mr. D. J. Page; (*c*) representing Local Groups: Mr. A. Stafford Beer, Mr. A. W. S. Clarke, Mr. M. G. Dixon, Mr. J. C. Eaton, Mr. D. A. Evans, Mr. D. A. Hopkins, Mr. W. N. Jessop, Mr. W. W. Jones, Mr. B. J. A. Martin, Mr. R. L. Plackett, Mr. R. B. Rowson, Mr. J. H. Simcock, Mr. C. S. Smith, Mr. A. F. Tout, Mr. J. R. Williams.

The Section Committee elected Miss J. Keen Chairman and Mr. M. Whyte Honorary Secretary.

The Officers of the Local Groups were:

Group	*Chairman*	*Honorary Secretary*
Birmingham and District	Mr. B. J. A. Martin	Mr. W. N. Jessop
Glasgow	Mr. W. Anderson	Mr. A. Kirk
Leicester	Mr. C. S. Smith	Mr. P. M. Hooper
London	Mr. A. F. Tout	Mr. A. W. S. Clarke
Merseyside	Mr. R. L. Plackett	Mr. J. R. Williams
North-Eastern	Mr. D. Gold	Mr. M. G. Dixon
Sheffield	Mr. A. Stafford Beer	Mr. D. A. Hopkins
South Wales	Mr. R. B. Rowson	Mr. W. W. Jones

Group membership on December 31st, 1955, was as follows (comparative figures for 1954 are shown in brackets):

Group	*Fellows*	*Non-Fellows*	*Total*	*Average Attendance*
Birmingham	42 (40)	21 (20)	63 (60)	25 (34)
Glasgow	8 (8)	32 (34)	40 (42)	20 (13)
Leicester	10 (10)	31 (31)	41 (41)	18 (19)
London	135 (121)	74 (74)	209 (195)	50 (51)
Merseyside	17 (13)	27 (35)	44 (48)	17 (17)
North-Eastern	28 (29)	7 (6)	35 (35)	23 (15)
Sheffield	34 (23)	27 (30)	61 (53)	17 (17)
South Wales	20 (18)	11 (9)	31 (27)	21 (18)
	294 (262)	230 (239)	524 (501)	— (—)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Fellows who are recorded as members of the Section are those who wish to be attached to a particular Group for the purpose of voting in its elections and being eligible to serve on its Committee.

The accounts of the Section and its Local Groups for the calendar year 1955, excluding expenditure by the Society in relation to the work of the Section, are summarized below:

Expenditure	£ s. d.	£ s. d.
Birmingham	22 3 6	
Glasgow	26 9 9	
Leicester	6 18 10	
London	22 4 3	
Merseyside	10 13 9	
North-Eastern	25 4 3	
Sheffield	18 15 1	
South Wales	8 18 6	
Section Committee	48 9 4	
		189 17 3
Carried forward to 1956		115 0 4
		£304 17 7

Income	£ s. d.
Brought forward from 1954	86 17 7
Grants from Council	218 0 0
	£304 17 7

Two one-day Conferences sponsored by Local Groups of the Section have been held during the Session. The first, held at Leicester in February, 1956, was entitled "The Third Eye of Management" and covered generally the field of statistical methods in industry. The second at Birmingham in May dealt with Statistical Methods and Industrial Experiments. Such Conferences are a valuable means of stimulating interest in statistical methods in industry, and the Council much appreciates the efforts made by the Local Group Committees in organizing them.

Fifty-one meetings, listed below, were held by local Groups: 7 each in Birmingham, Glasgow, London, and Liverpool; 6 each in Sheffield and Newcastle-upon-Tyne; 5 in Leicester and 3 each in Newport and Cardiff.

Birmingham and District Group

1955	
October 5th	The Uses of Factor Analysis in Industrial Physics—G. W. Scott Blair.
October 26th	A Method of Labour Turnover Analysis—K. F. Lane and J. E. Andrew.
November 3rd	The Role of Electronic Computers in Statistics—K. D. Tocher.
1956	
January 4th	Seasonal Variation in Time Series—J. R. N. Stone.
February 1st	Problems of Optimum Production Run—T. E. Easterfield.
March 7th	Statistical Tests—Do the Assumptions Matter?—G. E. P. Box.
April 4th	A Short Review of Some New Techniques—C. H. Leigh-Dugmore.

Glasgow Group

1955	
October 19th	Personnel Selection—P. A. D. Gardner.
November 24th	The Use of Statistics in the Study of Market Developments—J. V. Gregg.
December 14th	Visit to a local Textile factory.
1956	
January 25th	Simplified Statistics in Industry—John S. Ogilvie.
February 23rd	Control of Stocks—John Hepburn.
March 21st	Principles of Quality Control—A. Kirk.
April 26th	Techniques for Tomorrow—A film shown by the Ford Motor Company Ltd.

Leicester Group

1955	
October 20th	The Importance of the Elementary—B. P. Dudding.
November 10th	Upper Leather Wastage Allowances—K. W. Hall and J. V. A. Long.
December 1st	Statistics and the Accountant—W. F. Harris.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

1956

January 23rd — Criminology—L. T. Wilkins.
March 29th — Getting Control from Control Charts—M. J. Moroney.

London Group

1955

October 7th — Quality Control in the Paper Industry—I. F. Hendry.
November 4th — Clerical Operations as Statistical Processes—R. A. Fairthorne.
December 2nd — Linear Programming in Practice—E. Kay and W. E. Duckworth.

1956

February 3rd — Shortened Life Testing—G. A. Barnard.
March 2nd — The Handling of Uncontrolled Data—G. H. Jowett.
April 6th — Microspores and the Identification of Coal Seams—R. C. Tomlinson.
May 4th — The Sampling of Road Traffic—J. C. Tanner.

Merseyside Group

1955

October 4th — Applications of Electronic Computing in Industry—E. K. Robertson.
November 1st — Labour Wastage in Industry—H. Silcock.
December 6th — Monte Carlo Methods and Industrial Problems—W. N. Jessop.

1956

January 3rd — Ideas for Managers: Probability and Policy—M. J. Moroney.
February 7th — Some Problems in Consumer Testing of Convenience Goods—N. J. Squirrell.
March 27th — The Contribution of Statistics to the Study of Road Accident Causation—F. Garwood.
April 10th — A Discussion of Sequential Methods—F. Downton.

North-Eastern Group

1955

October 12th — Snap-Observation Techniques—K. T. Boyd.
November 9th — Some Applications of Statistics to the Improvement of Colliery Operations—B. H. P. Rivett.
December 7th — Elementary Principles of Survey Design—G. Park.

1956

January 25th — Some Aspects of Personnel Selection—W. Sluckin.
February 8th — Statistical Methods in Some Hospital Administration Problems—D. J. Newell.
March 14th — Graphical Testing of Binomial Frequencies—D. A. Evans.

Sheffield Group

1955

September 29th — Ultimate Capacity of a Slabbing Mill—H. G. Jones.
October 27th — Statistics without Computers—H. R. W. Watkins.
November 24th — Consumer Preference Testing—N. J. Squirrell.

1956

February 23rd — Stochastic Processes—D. A. Hopkins.
March 29th — The Theory of Games—D. V. Lindley.
April 26th — Figures for the Manager—A. Stafford Beer.

South Wales Group

1955

October 6th — Opinion Surveys—Mark Abrams.
November 23rd — Some Facts concerning Road Accidents in Great Britain—F. Garwood.
December 8th — Inventory Management—K. S. Lomax.

1956

January 27th — Some Aspects of Coding Material for Punched Card Analysis—J. A. H. Waterhouse.
February 24th — The Analysis of Continuous Measurement—K. W. Kemp.
March 21st — Description and Demonstration of a Periodograph—K. W. Kemp.
Accident Prevention Competitions—A. M. Whitehouse.
Three films: Frequency Curves; Techniques of Sampling; Marketing.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Council records its thanks to the Electric Lamp Manufacturers' Association, the Birmingham Exchange and Engineering Centre, Cardiff University, the Institution of Engineers and Shipbuilders in Scotland, King's College, Newcastle upon Tyne, the Leicester College of Technology, the Royal Institution in Liverpool, the Newport Technical College, and the South Wales Electricity Board for kindly providing accommodation for meetings.

The Study Section

The Section Committee for the Session has been: Mr. C. A. Moser (Chairman), Miss S. V. Cunliffe (Honorary Secretary), Mr. R. J. Finlayson, Mr. H. W. Gearing, Mr. D. L. Lamberth, Mr. E. Shankleman, and Mr. L. T. Wilkins. Mr. Moser and Mr. Wilkins were Council representatives. Mr. G. M. Clarke attended as the representative of the Bristol Group.

Meetings in London have been well attended both by members and visitors. The subjects chosen for papers have deliberately covered a wide range and the members of the audience have varied markedly with the different subjects. Attendances have averaged 53 (range 20 to 117).

The following papers were presented:

1955

October 12th	A Research Design for Studying the Effects of Television—W. A. Belson.
November 9th	The Scope and Limitations of Econometrics—L. R. Klein.
December 14th	Labour Wastage in Industry—H. Silcock.

1956

January 11th	Local Population Projections in England and Wales—J. R. L. Schneider.
February 8th	Advances in Recent Market Research Methodology—D. L. Lamberth.
March 15th	How the Statistician is helped by Electronic Computers—M. J. R. Healy, H. W. Durant, and S. J. Prais.
April 11th	Some Problems of Social Accounting (Lessons from a Study of the National Health Service)—B. Abel-Smith.
May 9th	Basic Statistical Concepts Illustrated on Tests for Randomness, with Special Reference to the Classical Problem of Points on a Line—F. N. David.
June 13th	Can Economists Forecast?—E. Shankleman.

The Committee of the Bristol Group for the Session has been: Mr. H. C. Mackenzie (Chairman), Mr. G. M. Clarke (Honorary Secretary), Mr. D. A. Clark, Mr. C. R. Hobbs, and Mr. W. Whitehead.

Six meetings have been held, and the following papers read and discussed:

1955

October 20th	Personnel Statistics—H. E. Bishop.
November 17th	The Application of Statistical Methods to Market Research Data—E. S. Holdsworth.
December 15th	Statistics in Hospital Planning and Design—N. T. J. Bailey.

1956

January 19th	The Dual Problem in Linear Programming—H. C. Mackenzie.
February 16th	Total and Difference in Samples of Two: Examples of Practical Statistics—M. J. Moroney.
March 15th	The Propensity to Save in Britain—L. R. Klein.

Attendances (ranging from 16 to 29) have averaged 23.

The Section has again benefited by the facilities provided for its meetings by the Electric Lamp Manufacturers' Association and the University of Bristol, to both of whom the Council expresses its gratitude.

The Medical Section

It was with much satisfaction that the Council in July, 1955, approved a proposal that the Study Circle on Medical Statistics, formed in 1951 as a group within the Study Section, be recognized by its constitution as a separate Medical Section.

The regulations for the conduct of the Section were approved in October, 1955, and are published in the *Journal*, Series A, Part 1, 1956.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Committee of the Section during the current Session was as follows: Dr. W. P. D. Logan (Chairman), Mr. M. P. Curwen (Honorary Secretary), Mr. B. Benjamin, Mr. B. B. Bonner, Dr. W. Hartston, Mr. J. A. Heady, Mr. P. D. Oldham, and Dr. D. D. Reid.

The following meetings were held at the Westminster Medical School at which attendances ranged from 17 to 47.

1955

September 27th — A Study of the Physiological Effects of Hot Climates—J. O. Irwin.
October 25th — Family Sickness as Reported by the Housewife—Lilli Stein and S. Sklaroff.
November 29th — From Graunt to Geneva: A History of Disease Classification—A. H. T. Robb-Smith.

1956

January 31st — Physique—Its Measurement and Influence (Joint Meeting with the Study Section)—R. W. Parnell.
February 28th — Ageing in Industry—F. le Gros Clark.
March 27th — Statistical Problems of Surgical Follow-Up—J. A. H. Waterhouse.
April 24th — Some Statistical Problems in Epidemiology—C. C. Spicer.
May 29th — Should Research be based on Routine Hospital Records?—A. C. Dornhorst, J. R. Bignall, and M. P. Curwen.

The Council again wishes to express its thanks to the Dean of the Westminster Medical School for kindly providing accommodation for the meetings.

The Society's Examinations

Twenty-two candidates for all or part of the Certificate and one for the Diploma presented themselves for the examinations held on March 21st, 22nd and 23rd, 1956.

Certificates have been awarded to the following Fellows:

Douglas, John Lunn.
Fletcher, Allan.
Gent, Michael.
Grimshaw, Jack James.
Holgate, Margaret Eleanor.
Holroyd, Edward Michael.
Lewis, Brian Davy.
Mawson, Peter Alan.
Waugh, William Anthony O'Neill.
Weaver, John Ernest.
Weekley, Edward.

The Council again records its thanks to the London School of Hygiene and Tropical Medicine for allowing the examinations to be held in its Department of Medical Statistics and Epidemiology, to those Fellows who acted as invigilators, and to the Examinations Committee and others who have helped the Committee in its work. The Committee has consisted of: Mr. B. Benjamin (Chairman), Professor R. G. D. Allen, Mr. N. T. J. Bailey, Professor G. A. Barnard, Dr. G. E. P. Box, Mr. W. J. Corlett, Mr. J. Durbin, Mr. E. C. Fieller, Mr. M. J. R. Healy, Miss R. J. Maurice, Mr. F. A. A. Menzler, Mr. A. Stuart, and Mr. A. M. Walker.

The Journals Series A and B

The four parts of the *Journal* Series A were published during the calendar year, and two parts of Series B. The Council again expresses its appreciation of the services given by Dr. J. O. Irwin as Chief Editor of Series B, and of the help given by the Editorial Panel—Mr. F. J. Anscombe, Dr. P. Armitage, Mr. N. T. J. Bailey, Professor G. A. Barnard, Professor M. S. Bartlett, Dr. D. R. Cox, Dr. H. E. Daniels, Dr. O. L. Davies, Mr. E. C. Fieller, Mr. J. M. Hammersley, Mr. D. G. Kendall, Professor M. G. Kendall, Mr. D. V. Lindley, Mr. R. L. Plackett, Dr. C. A. B. Smith, Professor J. R. N. Stone, Dr. K. D. Tocher, Dr. S. Vajda, and Dr. F. Yates.

It was with difficulty that the size of the *Journal* Series A was held in 1955 at almost the same number of pages as in 1954. From rather less than 500 in 1953, the number rose to 516 in the following year and by one further page in 1955. The 1955 volume of Series B contained 280 pages compared with 301 in the previous year, and 272 in 1953. As already mentioned, the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

printers notified the Society of a substantial rise in printing costs from the beginning of 1956 and that a further rise during the year must be expected. It is thus imperative that the Editors avoid any increase in the size of the *Journals*; indeed strenuous efforts must be made to reduce them to the size of the 1953 volumes or even less. Readers of papers and other authors are therefore urged to reduce their contributions to the absolute minimum. This point could be elaborated. but the Council feels that, if only by way of example, enough has been said.

The nineteenth and last of the second series of articles on the sources and nature of statistics in given fields, issued under the direction of Professor Kendall, is now in the press. A few more articles in the series had been contemplated, but it was felt that to wait for these would unduly protract the enterprise and 19 are sufficient to complete the second volume. The publication of this is now under discussion with Messrs. Oliver & Boyd who published the first volume.

Applied Statistics

Mr. Tippett and Dr. Beech have continued as joint Editors of *Applied Statistics* and the Council acknowledges its indebtedness to them for their services. The other members of the Committee for *Applied Statistics* have been: Dr. O. L. Davies, Mr. R. F. Fowler, Dr. J. O. Irwin, Mr. G. H. Jowett, Mr. C. A. Moser, Dr. C. Nuttall, Mr. E. D. van Rest, Mr. E. Shankleman, Mr. L. T. Wilkins, and the late Mr. Philip Lyle.

The completed volume for 1955 contained 212 pages. The Editors report that during the earlier part of the year suitable articles were scarce and the June issue was smaller than usual. The supply of contributions subsequently improved and it is hoped to maintain an average size of 72 pages a number. The articles have covered a wide range of subjects and have sought to appeal to readers at different levels of experience. They have, however, tended to be too long and there is a need for good short contributions that do not demand too much of the reader.

The financial surplus for the year, allowing for expenses additional to those in the publisher's account, was £186, compared with a profit of £166 in the previous year. While this result is encouraging it should be borne in mind that printing costs will rise.

The numbers of subscribers have been as follows:

	Fellows	*Non-Fellows*	*Total*
1953 . . .	490	912	1,402
1954 . . .	462	923	1,385
1955 . . .	467	992	1,459

The increase in the number of non-Fellows is gratifying and it is also good to see that the decline in the number of subscribing Fellows has not been repeated this year.

Library

Two projects approved by the Library Committee before the removal to Bentinck Street have been carried well forward during the year. The first was the creation of a small select library of the more important and frequently used books, arranged in subject order. By the end of 1955 this selection contained more than 1,000 volumes, classified according to the "Bliss" classification scheme. While none of the schemes considered appears to be entirely satisfactory for the Society's collection of books, this system, adopted experimentally, has so far proved acceptable, useful and convenient to Fellows using the Select Library. It will be modified or expanded in detail as experience provides knowledge of its adaptability to the Library's development.

The other project was to create a special collection of early statistical works, drawn mainly from books received through the Yule Bequest and from the many valuable and carefully chosen donations made to the Library by Mr. Yule since 1906. Such a selection, though not yet complete, has been made during the year and the chosen works have been arranged in chronological order and catalogued in some detail. Already Fellows interested in the early history and development of statistical literature are making use of the collection and finding it and its arrangement valuable.

An anonymous donor has presented for the Council Room a glass-fronted book-case harmonizing with the other furniture, this gift being a memorial to a late Fellow and holder of the Guy Medal in silver, Mr. George H. Wood. The "Yule Library" described in the preceding paragraph

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is appropriately housed in this bookcase. The collection and case together perpetuate the memory of two distinguished Fellows.

It was in 1855 that the Council, as stated in the report of the 21st anniversary meeting, "encouraged by the satisfactory state of the Society's finances . . . sanctioned an expenditure of £45 for the preparation of a catalogue of the books in its Library". The "distinct works" then numbered 2,000, most of which could be described as dealing with economics, social and demographic problems. Early editions of works by Graunt, Petty, and Richard Price were included as well as 43 items under A. Quetelet. The catalogue, beautifully hand-written and handsomely bound, has been displayed during its centenary year. It was compiled by B. R. Wheatley, well known later for his bibliographical work.

The following table presents an analysis of requests made for books by Fellows, during 1953, 1954, and 1955. The analysis does not include books consulted for reference in the reading rooms, whether by Fellows or visitors as this cannot be measured accurately. The Library was closed while the removal was in progress in 1953 and 1954.

	1953	1954	1955
Number of Fellows who made requests to borrow books	343	339	380
Number of applications made	1,319	1,032	1,370
Number of volumes borrowed	2,110	1,509	2,339
Number of volumes "on loan" (unavailable) at time of request	185	181	292
Number of volumes not possessed by the Library at time of request	174	137	266
Total number of volumes asked for	2,469	1,827	2,897
Non-serial works added to the Library	516	695	930

Finance

Abstracts of the Honorary Treasurer's Accounts, viz., the Statement of Income and Expenditure for the year 1955 and the Balance Sheet as at December 31st, 1955, together with the Auditor's Report thereon, are contained in Appendices B, C and D.

Total expenditure fell further from £13,633 in 1954 to £12,985 in 1955. Publication and distribution expenses accounted for £6,487 compared with £5,646 in the previous year, the increase being due mainly to the cost of the reprinting of two issues of *Journal* B and to the publication of a new Fellows List and of an index to *Journals* A and B covering the years 1940 to 1953. Salaries and Wages were slightly higher at £2,936 and there were small increases in Stationery and Miscellaneous Printing, and Library expenses. There were substantial reductions in both net rents payable and house expenses, and in Office and Library Equipment. The cost of alterations required to comply with new fire regulations and chargeable in 1955 amounted to £277.

Total Income, excluding Life Composition Fees received during the year, increased from £13,777 to £14,357. Annual subscriptions rose further by £249 to £6,650. Receipts from sales of *Journals* A and B and reprints, which declined from £4,606 in 1953 to £4,134 in 1954 rose to £4,677 in 1955, and the excess of publication and distribution expenses over receipts fell from £1,512 to £1,329. The surplus on *Applied Statistics* increased from £166 to £186.

The Rockefeller Foundation very generously agreed to extend for twelve months the period during which approved expenditure could be met out of the grant of $20,000 made in 1950. This made it possible for the Society to use the whole of the unspent balance, amounting to £1,115.

As a result of the year's operations there was a surplus of £1,372 as compared with £144 in 1954.

Composition Fees received during the year rose from £165 to £242. After allowing for the fees of compounders who died during the year or were presumed to have died in previous years, the Composition Fee Fund, maintained at the total of Composition Fees received from Fellows still living, amounted to £7,888 at the end of the year.

There was no further capital expenditure in 1955 to be met out of the Benefaction Fund. In the Balance Sheet the Society's Leasehold Premises, part of the 1950 Benefaction Fund, stand at £19,807, being the cost of £21,368 less amortization of £1,561. The Sinking Fund investment was increased by £425 (charged to Expenditure) to £1,561.

The Accumulated Fund of the Society increased from £8,786 to £10,158.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Council and Officers

Mr. Benjamin was appointed joint Honorary Secretary in place of Mr. Lyle, and Dr. Daniels was appointed to fill the vacancy on the Council.

The Council is pleased to report that Professor E. S. Pearson is able and willing to serve as President for a second year, and is glad to nominate him for that office during the ensuing Session.

The ordinary members of the Council who retire at the close of the current Session are Professor R. G. D. Allen, Dr. W. P. D. Logan, Mr. E. D. van Rest, Sir George Maddex, Mr. L. T. Wilkins, and Dr. F. Yates. Their services are acknowledged with gratitude. After observing the requirements of the Bye-Laws and considering the suggestions received, the Council recommended that the six vacancies for ordinary members be filled by the election of Dr. D. G. Beech, Dr. G. E. P. Box, Mr. M. P. Curwen, Dr. J. O. Irwin, Mr. F. A. A. Menzler, and Dr. C. A. B. Smith. No alternative proposals were received and the Fellows named below will be announced at the Annual General Meeting on June 20th, 1956, as having been elected President, other officers and other members of Council for the Session 1956–57.

Those whose names are marked ** were not Members of Council during the preceding session; those marked * have never served on the Council.

President

Professor E. S. Pearson, C.B.E.

Council

P. Armitage.
G. A. Barnard.
M. S. Bartlett.
*D. G. Beech.
*G. E. P. Box.
H. Campion.
D. G. Champernowne.
Stella V. Cunliffe.
*M. P. Curwen.
H. E. Daniels.
E. C. Fieller.
D. J. Finney.
R. G. Glenday.
D. Heron.
**J. O. Irwin.
Joan E. Keen.
M. G. Kendall.
**F. A. A. Menzler.
C. A. Moser.
W. F. Searle.
G. W. Sears.
Sir Robert Shone.
*C. A. B. Smith.
J. Stafford.
J. R. N. Stone.
L. H. C. Tippett.

Honorary Treasurer

R. F. Fowler.

Honorary Secretaries

R. F. George. Iris Douglas. B. Benjamin.

On behalf of the Council,
E. S. PEARSON, *President.*

R. F. GEORGE
IRIS DOUGLAS
B. BENJAMIN
} *Honorary Secretaries.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX A

From June, 1955 to May, 1956 inclusive, the candidates named below were elected Fellows of the Society:—

Abbas, Meher.
Acton, Rex Alan.
Aczel, Michael Aloysius.
Airth, Alan Don.
Anis, Abdel Azim.
Ardron, Geoffrey Henry Clarebrough.
Armitage, Peter Hutchison.
Ayoub, Suliman Taha.

Babik, Stanislaw.
Bacon, Francis William.
Bacon, James.
Baker, Victor Sydney.
Ball, Harry.
Banerjee, Keshab Chandra.
Bartlett, Brian Oliver.
Batkin, Richard Michael.
Benson, Richard Stamp.
Bessell, James Edward.
Biggers, John Dennis.
Bishnoi, Madan Lal.
Blackwell, David Harold.
Blatcher, Janet Thelma.
Bond, Raymond Charles.
Bonham, Donald Frederick Hurst.
Bransden, Bernard Walter Douglas.
Brodrick, John.
Bronstein, Norman Nathan.
Bruce, Victor Edward.

Cameron, James William.
Caverhill, Richard Stephenson.
Chan, Kum Chee.
Chavasse, Philip Ernest.
Chuck, Lesley Mary.
Clark, Arthur Wellesley Stewart.
Clegg, Dorothy Gertrude.
Clowes, Richard.
Coggins, Josephine Elizabeth.
Cotton, Charles Alan.
Cottrell, Bryce Arthur Murray.
Crisp, Donald Edward.
Crofts, John White.
Cullen, Charles Leslie.
Curnow, Robert Nicholas.

Dar, George.
Davies, Hilda Margaret.
Davison, Leonard Ernest.
Dick, Sylvia.
Dickins, John Cecil.
Doepel, Geoffrey Charles.
Donaghy, Thomas Stephen.
Drion, Edward Frans.
Duff, Hunter.
Dulley, Ronald Victor.
Duncan-Jones, Paul Daniel Picton.

Edwards, Hywel Wyn.
Eldridge, John Barron.
Ellinger, Alexander Geoffrey.
Evans, Iowerth Lloyd.

Farnfield, John Henry.
Figgins, James Edward Sullivan.
Findlay, Alan Frederick.
Fleming, Ralph Alexis Vernon.
Fletcher, Allan.
Flores, Ana Maria.
Flowerdew, Leonard James.

Gabriel, Kuno Ruben.
Gale, Alexander John.
Gani, Joseph Mark.
Gent, Michael.
Gilder, Robert Charles.
Godambe, Vidyadhar P.
Gower, John Clifford.
Granger, Clive William.
Green, Cyril Arthur.
Grimshaw, Jack James.
Gupta, Pramod Kumar.

Hagatong, Armando Florêncio de Oliveira.
Hall, Leonard Graham.
Hastie, Robert Allen.
Healey, Albert Thomas.
Heard, Frederick Richard.
Heath, Ernest Alfred John.
Helps, Sidney Martin.
Hockridge, Joan.
Holgate, Margaret Eleanor.
Holroyd, Edward Michael.
Hopkins, David Alan.
Horsfall, Anthony Victor Brian.
Horsnail, Alan Geoffrey.
Hughes, Ronald Owen.
Hyett, George Peter.

Ibbetson, David Noel Isserlis.

Jackson, Frederick Henry.
Jackson, John Stanley.
Jamieson, Andrew Thomas.
Jeffers, John Norman Richard.
Jeffrey, Alan.
Jenkins, Edward Leslie.
Jones, Alcwyn.
Joseph, Arthur Wolfe.

Kaitz, Hyman B.
Karpinos, Bernard D.
Kendall, George Roland.
Kerry, Clara Rose.
Knight, Gerald James.
Kruskal, William Henry.

Layton, Edgar Charles.
Lewis, Brian Davy.
Lomnicki, Zbigniew Alexander.
Lowe, William Donald.
Lutz, Walter.

Ma, Ronald.
Macaulay, Ronald Ian Murray.
McDonald, John Corbett.
McGlennon, Jack Gentry.
Mallows, Colin Lingwood.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Matthews, Philip Martin.
Matuszewski, Tadeusz Ignacy.
Maxwell, Sidney Reuben.
Miller, Frank Herbert.
Mills, John Lawrence.
Milnes, Henry Leslie.
Milstone, Dennis Michael.
Morawski, Kazimierz.

Nag, Anil Chandra.
Neill, Stanley Francis Mundy.
Newton, Norman Ernest.
el Nomrossy, Malik Mahmoud.

Oram, Matthew Henry.
Owen, George Lloyd.

Pain, Richard Henry.
Palmer, Valerie Mavis.
Park, Geoffrey Thomas.
Parker, John Michael.
Pennycuick, Kenneth.
Playfer, John Michael.
Plymen, John.
Potter, Kenneth Arthur Senior.
Powdrill, Ernest Arthur.
Purdon, Andrew Park MacDougall.

Quan, Chiu Hock.
Quinlan, William Patrick.

Rabbani, Mohammed Ghulam.
Reed, Graham Frederick.
Revell, John Robert Stephen.
Richards, Edmund Augustus.
Richardson, Peter Hanlon.
Robinson, Frederick David.
Rose, Harold Bertram.
Rose, Stanley Jack.
Rowntree, John Anthony.
Royston, Erica.

Scott, Christopher Saville O'Donoghue.
Scott, Elizabeth Leonard.
Shine, David Bernard.
Short, Joan.
Shukla, Vishwa Shanker.
Simms, Alfred George.
Simpson, Howard Russell.
Smart, John Victor.
Smith, John Sidney.
Snowdon, Eric Lewis.
de Smith, Catherine Joan.
de Somogyi, Janusz K.
Spector, Jacob.
Spencer, John Bernard.
Stearman, Roebert Lyle.
Stewart, Colin Macdonald.
Stone, Mervyn.
Sykes, George Stanley.
Syski, Ryszard.

Tan, Peter Cheow Khoo.
Thionet, P.
Thomerson, Frederick William.
Thorneycroft, William Terence.
Thong Ko-Sine, John.
Titmus, Colin Osborne Derek.
Touchin, Herbert Roy.
Trasler, Gordon Blair.
Tsao, Peter Kaineng.
Tucker, Ernest John.
Turner, John Christopher.
Turner, Joseph Frederick.
Twist, Anthony Frederick.

Vamathevan, Swaminathan.
Vandyk, Raymond Albert.

Warner, Bernard Thomas.
Weaver, John Ernest.
Webster, Denys Stuart Milroy.
Welding, Peter Ian.
Williams, Edgar Lloyd.
Williamson, William Henry Rowe.
Woodcock, Clive.
Woodrow, Guy.
Woods, John Francis.

Zaremba, Stanislaw Krystyn.

Corporate Representatives

Farquhar, Arthur Graham Francis, *representing* The Central Electricity Authority.
Gent, Alfred Walter, *representing* Standard Telephones and Cables Ltd.
Miles, Tudor Picton, *representing* The Iron and Steel Board.
Norton, Gordon, *representing* ORbit (Operational Research) Ltd.
Shenfield, Arthur Asher, *representing* The Federation of British Industries.
Starr, Anthony George, *representing* The Ministry of Supply.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPEN

STATEMENT OF INCOME AND EXPENDITURE

1954		EXPENDITURE			*1955*	
£	£			£	£
	311	Rents payable *less* rent received			101
	735	House expenses			437
	2,880	Salaries and wages (including contribution to staff superannuation scheme)			2,936
	250	Pension			250
	42	Insurance			63
	876	Office and Library equipment			155
	225	*Furniture previously forming part of Benefaction Fund assets now written off*			—
	177	Postage and telephone			123
	479	Stationery and miscellaneous printing			504
		Publication and distribution expenses:			
3,907		Journal, Series A (General) and reprints	3,844		
1,739		Journal, Series B (Methodological) and reprints	2,162		
5,646			6,006		
		Journal Index	228		
		Fellows' List	253		
	5,646			6,487	
		Library:			
246		Books	300		
130		Binding	132		
	376			432	
	187	General Meetings—ordinary and annual		181	
	12	Council and committee travelling expenses		22	
		Expenses of sections:			
47		Research	40		
246		Industrial Applications	222		
38		Study	34		
	331			296	
	9	Guy Medals		55	
		Examination expenses:			
92		Printing, stationery, postage etc.	67		
89		Examiners' fees	93		
	181			160	
	37	Auditor's fee (1954)		53	
	33	Miscellaneous expenses		28	
	425	Amortisation of leasehold premises		425	
	400	*Costs of removal to new premises*		—	
	21	*Loss on sale of Investment*		—	
	—	Cost of alterations to comply with fire regulations		277	
	13,633			12,985	
	144	Surplus carried to Accumulated Fund		1,372	
	13,777			14,357	
	165	Amount carried to Life Composition Fee Fund		242	
	£13,942			£14,599	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

DIX B

FOR THE YEAR ENDED 31ST DECEMBER, 1955

1954 £	£	INCOME	1955 £	£
		Annual subscriptions:		
6,286		Fellowship	6,523	
8		Research Section	12	
92		Industrial Applications Section	93	
15		Study Section...	22	
——	*6,401*		——	6,650
	50	Contribution from Royal Economic Society		50
	758	Dividends and interest (gross)		802
		Sales of Journal:		
2,458		Journal, Series A (General) and reprints	2,646	
1,676		Journal, Series B (Methodological) and reprints	2,031	
——	*4,134*		——	4,677
	166	*Applied Statistics* — Surplus for year		186
	13	Sales of other publications		11
	151	Examination receipts: Fees and Sale of papers		180
	1,587	Grant from Rockefeller Foundation—expenditure appropriated thereto		1,115
	—	Surplus from Conference		33
	102	Transfer from Life Composition Fee Fund, being fees of compounders known to have died during year and of those presumed to have died in previous years		89
	415	1950 Benefaction Fund—interest on investment		384
	—	Refund of Schedule 'A' tax paid in previous years		180
——				——
	13,777			14,357
	165	Life Composition Fees received during year		242
——				——
	£13,942			£14,599
——				——

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPEN

BALANCE SHEET AT

FUNDS AND LIABILITIES

1954			1955	
£	£		£	£
		Accumulated Fund:		
8,642		Balance at 31st December, 1954	8,786	
144		*Add:* Balance per Statement of Income and Expenditure ...	1,372	
	8,786			10,158
		Life Composition Fee Fund:		
7,583		Balance at 31st December, 1954	7,646	
165		*Add:* Life Composition Fees received during year	242	
7,748			7,888	
102		*Less:* Transfer to Statement of Income and Expenditure of fees of compounders known to have died during year, and of those presumed to have died in previous years	89	
	7,646			7,799
		Building Fund (per contra):		
1,486		Balance at 31st December, 1954	1,542	
56		*Add:* Income for year	58	
	1,542			1,600
		1950 Benefaction Fund (per contra):		
	30,971	Balance at 31st December, 1954		30,971
		Liabilities and Income held in suspense:		
2,266		Sundry Creditors	3,190	
1,836		Amounts received in advance	1,663	
	4,102			4,853
	53,047			55,381
		Frances Wood Memorial Fund (per contra):		
588		Balance at 31st December, 1954	571	
23		*Add:* Income for year	23	
611			594	
40		*Less: Prize awarded in* 1954	—	
	571			594
	£53,618			£55,975

REPORT OF

I have examined the above balance sheet and annexed statement of income and expenditure
December, 1955, and of its income and expenditure for the year ended on that date according to

5, London Wall Buildings,
London, E.C.2.
14th May, 1956.

CC-0: In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

DIX C

31ST DECEMBER, 1955

ASSETS

1954 £	£		1955 £	£
18,483		*Investments*, per schedule attached ... ... (Market Value, 31/12/54 £20,375, 31/12/55 £17,221)		17,968
		Building Fund:		
1,542		Investment, per schedule attached ... ... (Market Value, 31/12/54 £1,461, 31/12/55 £1,418)		1,600
		1950 Benefaction Fund:		
		Leasehold Premises:		
21,368		At cost	21,368	
1,095		*Less:* Amortisation to 31st December, 1955	1,561	
20,273			19,807	
1,095		Sinking Fund Investment, per schedule attached (Market Value, 31/12/54 £1,138, 31/12/55 £1,408)	1,561	
21,368			21,368	
9,554		Investment, per schedule attached ... ... (Market Value, 31/12/54 £10,791, 31/12/55 £9,095)	9,603	
49		*Interest accrued*	—	
	30,971			30,971
		Current Assets:		
426		Debtors and amounts paid in advance	873	
100		Arrears of subscriptions recoverable (estimated)	100	
177		Interest accrued on investments (gross)	231	
1,348		Balances at banks and cash in hand	3,638	
	2,051			4,842
	53,047			55,381
		Frances Wood Memorial Fund:		
300		Investment, per schedule attached ... ... (Market Value, 31/12/54 £449, 31/12/55 £377)	300	
271		Post Office Savings Bank Deposit	294	
	571			594
	£53,618			£55,975

NOTE.—

No value is placed in the Accounts on—

(a) Journals and other publications in stock.

(b) Books in library.

(c) Pictures, furniture and equipment.

E. S. PEARSON, *President.*

R. F. FOWLER, *Honorary Treasurer.*

THE AUDITOR

which, in my opinion, respectively give a true and fair view of the Society's affairs as at 31st the best of my information and the explanations given to me and as shown by the books.

A. RAE SMITH, *Chartered Accountant.*

(*Deloitte, Plender, Griffiths & Co.*)

Auditor.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX D

SCHEDULE OF INVESTMENTS—31st DECEMBER, 1955

1954			*1955*		
Market Value £	*Book Value* £		*Nominal* £	*Book Value* £	*Market Value* £
6,948	*5,580*	2½% Consols (Guy Bequest)	10,528	5,580	5,948
1,476	*1,185*	2½% Consols	2,237	1,185	1,264
1,634	*1,324*	3½% Conversion Loan	1,865	1,324	1,410
435	*490*	3½% War Loan	500	490	381
2,658	*2,992*	3% Savings Bonds, 1965–75...	2,970	2,992	2,328
2,000	*2,000*	3% Defence Bonds	2,000	2,000	2,000
170	*100*	British Transport 3% Guaranteed Stock, 1978–88	195	100	163
5,054	*4,812*	British Electricity 3½% Guaranteed Stock, 1976-79	4,535	4,297	3,727
£20,375	*£18,483*			£17,968	£17,221
		Building Fund:			
£1,461	*£1,542*	3½% Conversion Loan	1,723	£1,600	£1,418
		1950 Benefaction Sinking Fund:			
£1,138	*£1,095*	British Electricity 3½% Guaranteed Stock, 1976–79	1,713	£1,561	£1,408
		1950 Benefaction Fund:			
£10,791	*£9,554*	British Electricity 3½% Guaranteed Stock, 1976–79	11,006	£9,603	£9,095
		Frances Wood Memorial Fund:			
£449	*£300*	British Transport 3% Guaranteed Stock, 1978–88	513	£300	£377

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PROCEEDINGS OF THE ONE HUNDRED AND TWENTY-SECOND ANNUAL GENERAL MEETING OF THE ROYAL STATISTICAL SOCIETY held at the London School of Hygiene and Tropical Medicine on Wednesday, June 20th, 1956, at 5.30 p.m.

The Chair was taken by the President, Professor E. S. Pearson, C.B.E.

The President presented the Report of the Council for the financial year 1955 and the session 1955–56, and moved that it be received.

During his submission of the Report the President presented Certificates to the successful candidates in the 1956 examination who were able to be present at the meeting.

The Honorary Treasurer, Mr. R. F. Fowler, presented the accounts and seconded the motion that the Report be received.

On being put to the meeting the Report was formally received.

On the proposal of Professor Bradford Hill, seconded by Mr. Menzler, Sir Alan Rae Smith was unanimously re-elected Auditor of the Society for the session 1956–57, at a fee of fifty guineas.

The President announced that as no alternative nominations had been received the President, Council and Officers for the session 1956–57, nominated as shown on the list already circulated, were duly elected.

The meeting then terminated.

The Annual General Meeting was preceded by an Ordinary Meeting at which the following were elected Honorary Fellows of the Society:

Professor Oskar Anderson	Germany.
Professor Ragnar Frisch	Norway.
Professor Harold Hotelling	United States.
Professor Charles Jordan	Hungary.

The candidates named below were elected Fellows of the Society:

Victor Charles Ablewhite
Stanley Harold Baker
Peter James Bannister
Anthony Michael Beaumont-Dark
Richard John Briggs
Donald McGregor Carter
Richard Leston Carter
Victor Edward Firmin
Donald Alexander Stuart Fraser
Walter Bruce Hall
Claude Herbert Hitching
Eric Selby Long
Arthur Alfred Munday
Neil William Please
Derek Salman Pugh
Nathan Stephen Roseman
Anthony Peyton Vincent

Corporate Representative

George Anthony Hepworth, *representing* Hepworth & Grandage, Ltd.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REVIEWS OF STATISTICAL AND ECONOMIC BOOKS

CONTENTS

PAGE

1.—*Devons* (*E.*). An Introduction to British Economic Statistics 478
2.—*Wallis* (*W. Allen*) & *Roberts* (*Harry V.*). Statistics: A New Approach . . 479
3.—*Cornell* (*F. G.*). The Essentials of Educational Statistics 480
4.—*Grünhut* (*M.*). Juvenile Offenders before the Courts. 480
5.—*Hickman* (*C. Addison*) & *Kuhn* (*Manford H.*). Individuals, Groups, and Economic Behavior 481
6.—*Morton* (*J. E.*). Urban Mortgage Lending 482
7.—National Bureau of Economic Research. Capital Formation and Economic Growth 484
8.—World Health Organisation. Annual Epidemiological and Vital Statistics . . 484
9.—*Barker* (*R. E.*). Books for all 486
10.—*Roy* (*René*) Ed. Cahiers du Seminaire d'Econometrie No. 3 Les Modèles Économétriques 486

1.—*An Introduction to British Economic Statistics.* By E. Devons. Cambridge University Press, 1956. vii, 255 pp. 7¼″. 22*s.* 6*d.*

This book is described as being designed principally for University students of economics and commerce. Those who have attempted to teach students about the sources of economic statistics will realize the great value of a book giving information about the most important of these statistics. The non-specialist student cannot be expected to read all the White Papers, articles, etc., many of them too technical for his needs, which it is otherwise necessary to suggest as references. Devons has done a very thorough piece of work in providing the necessary summary for these purposes. He also gives plentiful references to aid those who want more detailed information than he is able to give on any of the topics discussed.

One unfortunate feature of a book of this nature is that it inevitably becomes rapidly out of date. Even by the date of publication, there have been several changes in the information available from that which existed when the book was written. Publication of several new statistical series has started, some of the existing series have been revised, as in the case of the index of retail prices, and increased information is available about others, as, for example, national income. Such difficulties are inevitable but they do call for a fairly frequent revision of a book of this nature. Even when the discussion no longer contains the most recent information on any topic, the book will remain a useful source of reference for those interested in using past statistics, but, for the economist interested in current economic statistics, much of the value of the book will be lost. For both types of user, it might have been helpful if Devons had stated rather more clearly the date up to which he has written.

It is possible to make a few minor criticisms of the presentation, where the student may be slightly misled or left a little baffled. The attempt to show when changes of weights for an index of industrial production matter and when they do not is very confusing. Devons tries to make the point that changes in the relative sizes of the different industries do not matter in themselves. What matters is whether the ratio of net output to the volume of output changes in the same proportion in the different industries or not. This is put as "a change in the relative movement of net output and the volume of output in different industries". Realizing that the idea is difficult, Devons repeats it in almost the same words, but it seems almost impossible to put an unambiguous meaning to this sentence. The introduction of the undefined concept of relative net output prices only adds to the confusion.

The early chapter on Population is, perhaps, the most likely to cause difficulty. The introduction of the stationary population by assuming births in each year to be just sufficient to offset deaths seems less useful than an assumption of a constant number of births. It makes the attainment of constant total numbers immediate and not dependent on the properties of the mortality rates at all. The stability of numbers of each age and sex is also not achieved on Devons' assumption without some unstated (even if plausible) restrictions on the specific death rates. It is also doubtful whether the relationship between the average expectation of life and the life-table death rate will be as obvious to the student as he assumes. The description of expected deaths for the Registrar-General's comparison of deaths in different regions in terms of the life-table death rates for the country as a whole is also rather confusing—particularly as it immediately follows a description of comparisons of occupational mortality (in principle the same) in terms of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

specific mortality rates for the country as a whole. The footnote to the statement that internal migration could in principle be calculated from Census reports and natural increase for small regions and towns states that all these figures would be based on deaths occurring in the region. This is a little misleading when deaths are classified by normal residence although it is valid for the application given to service deaths overseas.

One final point appears a little baffling when compared with the published account. This is in the description of the Ministry of Labour's estimates of the number of insured employees in different industries. It is hard to understand from the official description of the use made of employers' returns how these would help to overcome any uncertainties about the reliability of the industrial classification. Any mistakes in classification would surely be left in the residual figure of cards for employers not making returns and so get carried through to the final estimates.

It is hard to find any important references which are not cited, although there might, perhaps, have been a reference to the articles by Nicholson and Penrice in the discussion of the relationship between the movements of wage rates and earnings. Rather surprisingly, Devons does not seem to have got help from official statisticians in clearing up some of the doubtful points of interpretation and is reduced on some occasions to presumptions about what is done.

The nature of the subject matter makes it difficult for a book like this to afford light and entertaining reading. In some ways it is a pity that many of the more interesting points are put in footnotes. The book might have been more palatable to the average student if it had been found possible to incorporate some of these points in the text. Despite this and despite the other minor defects, Devons has produced a book which admirably serves the purpose for which it was intended and which fills a gap which has existed for far too long. May we hope that this description of British economic statistics will be kept up to date by frequent revisions. W. J. CORLETT.

2.—*Statistics: A New Approach.* By W. Allen Wallis and Harry V. Roberts. Glencoe, Illinois. The Free Press. 1956. xxxviii, 646 pp. 9¼″. $6.00.

Most people, when they see the rather striking dust-cover of this book, remark "What is new about it?" The most obvious new feature is its length. When it is borne in mind that the book is only an introduction to Statistics, that it only develops the simpler techniques and avoids all proofs, the sheer weight of the volume is surprising. How has this come about? Principally because of a prolix style and the lavish use of examples, supplemented by the fact that the treatment is not confined to one field of application but discusses uses of statistics in many sciences. (Economic and social examples predominate.) The examples are nothing short of wonderful. They are almost all taken from cited texts, ranging from Thucydides on the use of the mode to the Chicago Tribune on the Drabik family of that city, which has seven daughters all called Mary. They are not merely striking examples, likely to arouse the interest of the dullest student, but are relevant and revealing, and must have entailed a lot of hard work in collection. Another new feature is the reduction of most tests of significance to the normal form so that a single table of that distribution suffices; even Student's t is treated in this manner, giving a procedure which is surely more complicated than the use of a new table. Procedures based on ranks are also frequently used, for example in the comparison of several means, and the resulting F-tests of normal theory avoided.

The book is divided into four parts entitled: The Nature of Statistics, Statistical Description, Statistical Inference and Special Topics. The first part (166 pages) describes the scope of statistics, illustrates its uses and misuses, introduces the ideas of sampling and concludes with a chapter on measurement and observation. Because of the illustrative material this section is particularly long; by page 100, before the introduction of sampling ideas, we have learnt no new ideas or had developed any new tools, and the best that can be said is that our eyes have been opened to the possibilities of the statistical method. The second part (142 pages) is along more usual lines including the art of organizing data and the measurement of averages, variability and association. The latter part of the first section and the whole of the second are excellent: the authors are doing little more than making common sense of numbers, but they do it with a competence which one can but envy. Any intelligent person could read this section and be better able to criticize and understand the world: it should be compulsory reading for all journalists. Probability ideas are introduced in the third part (166 pages) and sampling theory, principally in connection with the normal law, is explained. A long discussion of the two types of error introduces some test procedures and the section concludes with a chapter on estimation. The emphasis here is questionable: it would surely be better to reduce the material on Type II errors, which are rarely of use outside sampling inspection (and even there an honest use of losses and prior beliefs is more relevant), and increase the material on confidence intervals and efficiency of estimates. It is a curious feature that although the authors perform special feats to reduce significance tests to the normal

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

form, the remarkable properties of this law in possessing sufficient and efficient statistics is not discussed. The final part (148 pages) proceeds at a faster pace than the remainder and includes chapters on the design of investigations (experiments and surveys), quality control, relationships between variables, time-series and a final one on "short-cuts" dealing mainly with ranking methods and binomial probability paper. The authors are confident that "what is presented . . . will remain sound and useful half a century hence". The reviewer does not share this confidence. Tables of square roots, random numbers and the normal curve are included. The book is well indexed and besides a detailed list of contents there are lists of examples, tables and figures. The members of this trio have the same numbers as the pages on which they appear: a system of reference which is most convenient. Two examples on the same page are distinguished by letters, thus, 19A.

In passing comment on this book fair attention must be paid to the remarks in the authors' preface which, in tone with the rest of the book, is long, running to seven pages. The orientation is towards those who will not be specialists in statistics and who are virtually devoid of mathematics. It is surprising that such persons should have time to spend on this lengthy course, in a field which is not their speciality. (In common with most American books the material has been used in lecture courses.) The authors make the surprising claim that the "avoidance of mathematics is a virtue". This seems hard to understand when the basic ideas of the subject are essentially mathematical in character. To use this book for students with knowledge, for example, of the calculus, would be almost as bad as teaching a child to write without using his hands. A statistics student without mathematics will always be handicapped. It is only fair to mention in connection with the size of the book that the authors "expect to prepare almost immediately a manual for teachers which will indicate . . . the possibilities for picking and choosing among the topics".

To summarize then, this is a prolix book for the non-specialist with no mathematics, orientated more toward the social than the experimental scientist. Within these qualifications the book is well-written, the topics sensibly chosen in general, and mistakes are very rare. The examples are excellent and every teacher of statistics will welcome them. It is perhaps especially suitable for the student working on his own: there are sets of exercises, headed "Do It Yourself". Much of the first half can be warmly recommended for the general serious reader. The price, per ounce, is quite moderate. The printing is good. D. V. LINDLEY.

3.—*The Essentials of Educational Statistics.* By F. G. Cornell. New York, Wiley. London, Chapman & Hall, 1956. xii, 371 pp. 9″. 46*s.*

This is, in effect, a text-book on elementary statistics with numerous illustrations of applications to data obtained in the activities of the educational system of the United States. A low level of mathematical ability is assumed, amounting to little more than familiarity with the use of symbols and simple algebraic manipulation. The notion of an integral as representing the area under a curve is introduced, but techniques of integration are not discussed.

The usual elementary topics are introduced—classification, frequency distributions, central values, dispersion, skewness, kurtosis, distribution of mean, variance and correlation coefficient, chi square, comparison of means and proportions, partial and multiple correlation, and simple analysis of variance. There is also a discussion of "test reliability", defined as the correlation between the actual observed scores and the "true" scores of a test. It is pointed out that this cannot be measured directly since the "true" score cannot be observed, but it is stated that this difficulty can be overcome by applying two "parallel" forms of the same test. The assumption that these two forms would have the same "true" score does not appear to be questioned.

As in most modern text-books, the formal presentation of the subject-matter is clear and in accordance with fashionable notions on testing hypotheses, estimation etc., but there is no indication of possible alternative approaches and the student will probably gain the impression that the techniques described are somehow inherent in the natural order and therefore unassailable. An extension of the final section of the first chapter—on "Logic Commonsense and Statistics"—could quite well emphasize the necessity for critical reading.

The book could reasonably be recommended to beginning students in statistics, were the cost somewhat lower. The limited nature of the examples would not prevent the book being useful to students whose major interests lie in other fields. However, a judicious pruning of the text might be desirable, and would perhaps result in a lower purchase price. N. L. JOHNSON.

4.—*Juvenile Offenders Before the Courts.* By Max Grünhut. Oxford, Clarendon Press, 1956. vi, 143 pp. 8½″. 21*s.*

Knowledge of the work of the juvenile courts, particularly in relation to their social setting, has been one of the major gaps in the study of juvenile delinquency in this country since the war.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Dr. Grünhut deserves credit for pioneering in this field. As an empirical investigation, however, his study immediately throws into relief the problem of the non-statistical approach to a situation which can only be handled effectively by multivariate analysis. It must be said that the major shortcoming of this book is that it is almost entirely based on the use of statistical material, but gives the impression of a lack of trust in statistical methods. The predominant technique is analysis by inspection—the use even of correlation coefficients is largely rejected, no tests of statistical significance appear at any point, but there are a few, somewhat grudging, references to "statistical expectation". The book therefore is of virtually no methodological interest.

In spite of its archaic techniques, as the only study of the subject, the evidence and conclusions deserve some attention. The first section discusses the distribution of juvenile offences in terms of the police districts in England and Wales and relates these to certain indices of social conditions in general. The method used is to isolate areas of extreme deviance and to consider their rates of infant mortality, adult criminality, proportion of small incomes etc. This is most painstakingly done, but one must regret the author's reluctance to use covariance methods which, indeed, he pessimistically rejects: "It is too early to work out the prevalence of factors connected with crime and delinquency in different localities in terms of statistical correlations". (p. 45). Possibly the most interesting conclusion in this section is that which points out the relative stability of the pattern of delinquency, particularly in the areas with extremely high delinquency rates; Merseyside for example was the outstanding area for high juvenile delinquency rates both in 1936–8 and 1948–50.

The next section analyses the extent to which Courts in different areas use the various forms of treatment available to them. There is an extended and interesting discussion of the post-war tendency to impose fines in preference to placing offenders on probation. Here, the author's familiarity with court and police procedures adds considerable illumination to his account of the changes in practice.

The most valuable section is that which attempts to link treatment policy with characteristics of the offenders brought before the courts. This is done for five selected areas with a hundred cases sampled from each. Every juvenile delinquent in the sample was assessed in terms of a number of "criminological and social factors"; rates were then calculated for the presence of these factors both singly and in combination among the delinquents of the area. Results show that a high proportion of delinquents come from "good" homes, and it is demonstrated that committal to an Approved School is usually reserved for offenders from "bad" homes; the most concrete evidence of a "bad" home in this context is the presence of another delinquent in the family.

The other main conclusion in this section is that the use of absolute and conditional discharges as a treatment for second and later offenders is generally unwise because the chances of reconviction are high. The chief weakness of this section is that no reasons are stated for the selection of the variables used, nor are they defined. It appears, in fact, that no precise definitions were formulated; it is said that "in order to achieve the necessary uniformity, the assessment of the data in the case studies was in the hands of one research worker", but although this presumably helped to some extent, its value must have been limited by the fact that the case studies were themselves derived from many different sources—schools, probation officers, children's officers, police authorities etc. The difficulties of assessing such factors as "lack of discipline", "family disharmony", "low intelligence" on the basis of such subjective reports would seem considerable.

Dr. Grünhut's study, which is very modestly written, may be said to ask the right questions, to indicate some of the difficulties of answering them, to produce some interesting hypotheses, and to point the way for a more systematic, sophisticated and rigorous survey.

J. P. MARTIN.

5.—*Individuals, Groups, and Economic Behavior.* By C. Addison Hickman and Manford H. Kuhn. New York, Dryden Press, 1956. xvii, 266 pp. 8¼". $4.75.

The authors of this book seek to show how a knowledge of social psychology is necessary and helpful to an economist even at the theoretical level. A secondary object appears to be to assert the superior claims of a particular school of social psychologists. The ideal reviewer would be equally well versed in both the disciplines with which the discussion is concerned. As it is, however, the reader must be content with the impressions of a possessor of a smattering of economic knowledge on being told that a substantial dose of social psychology would be good for him.

In these circumstances it is desirable that the bias of the reviewer should be revealed at the outset. Great advantages undoubtedly stem from keeping economists up-to-date with advances in the branches of knowledge which impinge upon their own speciality. But if these steps forward are to be put to practical use in economics, it is necessary that the science in which they are made

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

should have reached approximately the same degree of organization and development as economics herself. Most economists are painfully aware of the primitive state of many parts of their own speciality, but it is probably more than a personal impression that her sister subject has still considerable leeway to make up. Under present conditions a closer integration of economics and social psychology seems to involve the exchange of over simple but precise generalizations for vagueness and a considerably expanded vocabulary of professional jargon. Economists ought to remain intelligently aware of progress but also somewhat wary of abandoning too readily their own well tested discipline.

The authors open with an illuminating synopsis of social psychological theory. Freudian, field and learning theory are all reviewed, but pride of place is given to "self" theory which is also called "reference group" theory. In the author's words "As self theory views the individual he derives his plans of action from the roles he plays and the statuses he occupies in the groups with which he feels identified—his reference groups". The theory is then applied to three economic problems: those of managerial motivation; of interpersonal comparisons of utility; and of the rival claims of planning and freedom. In each case the economics of the problem are described and the views, conflicting and otherwise, of distinguished economists are freely quoted. The heavy artillery of social psychology is then deployed to resolve the issue.

In the discussion of managerial motivation considerable energy is dissipated tilting at the windmill that economists believe that entrepreneurs and corporations maximize profits regardless of all other considerations. Even in the dream world where the future is certain, profit maximization is only deemed to be the aim if other things are equal. The fact that profit maximization is the basic premise of much abstract theorizing only reflects the recognition that it is extremely difficult to quantify the other things that do not remain equal or to specify the managerial reactions which are likely to result. Where, however, an uncertain environment is an essential feature of the problem, it is generally recognized that profit maximization has no obvious meaning.

The reference of this problem to social psychology involves a lengthy description of what the corps of managers in American Society think of themselves and of how they are regarded by the rest of society. The results of various attitude tests are extensively quoted. The upshot of the argument is that "The evidence presented indicates that contemporary corporation executives are characterized by self attitudes whose motivational goals include social prestige, social leadership, maintenance of their firms as going concerns, trusteeship in several other senses, and the persistent application of energy to their tasks".

The attack on the problem of interpersonal comparison of utility involves a demonstration that certain social groups are more deeply attached to particular possessions and amenities than others. For instance, graduates successfully resist the charms of television to a noticeably greater extent than the possessors of more humble academic qualifications with the same dollar income. Although its relevance to the theoretical problem under discussion is not altogether obvious, the authors comfort themselves with the conclusion that "we should ultimately find ways of pinpointing an individual in his reference groups so accurately that we can predict from the reference group values to the individual's pattern of satisfactions".

The planning and freedom controversy is conducted at a rather different level. At this point in the book an overshort acquaintance with social psychology became an overwhelming handicap to the reviewer. His struggles only yielded the unrewarding conclusion that the controversy could only be satisfactorily resolved if it were realized that different social groups meant different things by the words "planning" and "freedom" in this particular context.

The final section of the book deals with the research methods used in social psychology to determine and analyse the attitudes of individuals and social groups. Apparently one of the more powerful tools available consists of a "self-attitude" test in which the subject is asked to write down twenty different statements in response to the question "Who am I?"

The authors have performed a useful task in demonstrating what social psychology has to offer to economists. It is to be hoped that they will have many readers who can make profitable use of their gift and that this review has erred on the side of over-cynicism.

A. D. Roy.

6.—*Urban Mortgage Lending. Comparative Markets and Experiences.* By J.E. Morton. Princeton, University Press, 1956. London, Oxford University Press. xxi, 187 pp. 9″. $4.

In the United States, the National Bureau of Economic Research has previously sponsored a number of studies on the urban mortgage market. This volume, prepared by Dr. Morton, Professor of Statistics at Cornell University, summarises the findings of these earlier investigations, and brings together the material relating to the main Institutional Lenders (Insurance Companies, Commercial Banks, Mutual Savings Banks, and the Savings and Loan Associations).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Chapter I deals with the non-farm mortgage debt, tracing changes in this item over the 33 years from 1920 to 1953. Tables show the growth of such Debt year by year, distinguishing between corporate and private borrowing, and showing also the proportion of the debt secured on residential properties, commercial properties, etc. Debt-to-value ratios are examined at intervals throughout the period for various types of property, and for the different regions. Since 1934 Government sponsored home mortgage insurance has developed rapidly, until such loans represent some one-third of the total non-farm debt.

In Chapters II and III the author studies the structure of urban mortgage markets, analysing the share of the market held by private persons, savings and loan associations, commercial banks, individual savings banks and the life insurance companies. Over the whole period examined, the life insurance companies secured an increasingly important share of the business. For one to four family home mortgages, however, savings and loan associations have remained the most important source of finance. The share of individual mortgage investors has continued to decline.

Details are given of an interesting sample investigation, dealing with the characteristics of outstanding mortgages. Some 24 large life insurance offices, together with 170 commercial banks, and 202 savings and loan associations, were included in the sample; altogether some 11,000 mortgages, outstanding at mid-1947, were covered by the investigation. Unfortunately, the response to the inquiry by the smaller institutions was deficient, so that the sample estimates are subject to a non-response bias of unknown magnitude. From this sample, much interesting information is obtained regarding such details as the contract length, the loan-to-value ratios, and the interest rates, analysed according to the type of property and the category of lender. The Tables and Charts showing the long term trend of mortgage interest rates would, however, have been much more informative if accompanied by data as to the corresponding yields on long term Government and Corporate Bonds. Such information, readily obtainable from published yield indices, would reveal any trends in the basis for mortgage interest rates, free from fluctuations in general interest rates.

In an attempt to study the lending experience over the last 30 years, a further sample investigation was undertaken. The sample consisted of some 20,000 loans, made between 1920 and 1947. The 24 life insurance offices again contributed to the sample, together with 116 commercial banks and 92 savings and loan associations. The author has again to draw attention to the limitations of his data, due to the non-response bias. In particular, many small lending institutions suffered such severe losses during the 1929 depression that they were wound up; consequently their experience, which would have been particularly interesting, was not available. The sample study examines the foreclosure experience according to the type of lender, the class of property, the amount of loan, etc. A particularly interesting section of the investigation deals with the liquidation experience (i.e. the loss per cent of original loan amount), and with the realized yield, that is the yield obtained after all losses had been discharged. As regards foreclosure rates, etc., borrower characteristics and the social and economic aspects of the loan would seem of special importance. Unfortunately, the requisite data for this purpose are largely non-existent, apart from a small and specialised experience in respect of the Home Owners' Loan Corporation.

From the statistical point of view, the most interesting section of the book is to be found in the appendix, where a full and detailed explanation is given of the sampling methods used, and of the various difficulties that were encountered. As the objective was to compare the characteristics of current loans and the experience of past transactions for the various institutional lenders, it was necessary to include all classes of lender, large and small. In view of the difficulties due to non-response and lack of historical information among the smaller lenders, this section of the experience is of rather limited value. In practice, only the life insurance offices with their extensive records and clerical facilities were able to furnish reliable historical information. Presumably, if the investigation had been confined to this class of lender, a much more detailed and reliable analysis of mortgage experience would have been possible. It is surprising that consideration has not been given to the feasibility of a more detailed survey, confined to these larger lending institutions. Such a survey, conducted on present lines as regards foreclosure rates, liquidation experience, realised yield, etc. might be extended to compare the results according to such aspects as the borrowers occupation, the class of property etc. In this way, information of value to the conduct of future mortgage operations would be obtained.

In conclusion, the book is of distinct interest to students of statistical sampling techniques. It is of little value as a guide to mortgage policy in the U.K., as U.S. mortgage conditions and methods are so different from the practices in this country, and as no attempt is made to study the effect of the borrowers' social and economic conditions on the results of the loan transactions. Nevertheless this U.S. experience is instructive and worthy of study, revealing as it does the losses that can be sustained through mortgage operations in times of depression, and the margins that may

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

be required in interest rates to cover such losses. In this connection, the practice of U.K. lending institutions, granting mortgages on a large scale at interest rates only just above the gilt-edged level, is surely somewhat open to question. J. PLYMEN.

7.—*Capital Formation and Economic Growth. A Conference of the Universities-National Bureau Committee for Economic Research.* A Report of the National Bureau of Economic Research, New York. Princeton University Press (Oxford University Press), 1955. xiii, 667 pp. 9″. $12.00.

This substantial volume contains the papers presented at a conference which met in 1953 under the auspices of the Universities-National Bureau Committee for Economic Research. It also contains numerous invited and uninvited discussion papers and authors' rejoinders. The thirty contributors include an impressive proportion of the well-known American specialists in problems of economic development and they cover between them a wide range of viewpoints. The result is an exceptionally solid contribution to an important aspect of the problem of economic growth.

The papers are focused on three main topics—sources and channels of finance (there are two papers for the capitalist countries and two on the Soviet Union), the influence of enterprise and business organization (including two papers on developed and two on underdeveloped countries), and the influence of technological progress. In addition there is a theoretical paper by Lowe setting up an ideal model of the growth process, and two general reviews of the papers as a whole—one an introduction by Abramovitz and the other a summing-up by Rostow. The discussion consists of a series of prepared papers. In effect, including the two general reviews, papers written about the main papers take up more than a quarter of the whole volume.

Often the inclusion of discussion in the published results of a conference of this kind adds to the disjointed effect of the symposium. A series of brief comments on arbitrarily selected points can leave an irritating impression of superficiality, even when the points are worth raising and well taken. In this case, however, the procedure has been a success, largely because the discussion has been more considered than spontaneous. Not only have the contributors, as Rostow points out, been reading each other's articles and books: they also appear to have digested each other's papers to an extent which is unusual among conference participants.

Faced with such a variety of material, some of it rich in detail, and with so much discussion and counter-discussion, the reviewer is obliged to retreat even further than is usually necessary in order to take a general view of this volume. From a more detached vantage point three main related features seem to stand out as characteristic of its varied contributions to the long-term problems of capital formation.

The first is the strong empirical bias of the discussion. With the exception of Lowe's exercise in model-building, which stands avowedly apart from the other papers, each of the contributions is directly related to actual case histories or heavily reinforced by factual illustration. The scarcity of reliable information on long-term capital trends makes it necessary to base some of the factual illustrations on a highly subjective interpretation of the available data, but the empirical foundation is still the main support of the resulting analysis.

The second feature is the degree of emphasis laid on non-economic, and particularly on the institutional, factors in capital formation. It is more usual for economists to pay lip service to the sociological determinants of economic behaviour and then to treat them as invariable. In this collection, however, the economic historian is predominant and the sociologist has been given an important role.

The third feature is a marked predilection for international comparisons. The papers by Kuznets and Goldsmith set the stage with a series of common and uncommon exercises in comparative economic statistics. Some of the other papers accept a more restricted geographical horizon, but all of them implicitly or explicitly base an important part of their analysis on the significance of international similarities or differences.

Like most conference volumes this is not the kind of book to be read at a sitting, but it does achieve a larger measure of intellectual unity than most. In spite of its somewhat formidable bulk, it will be valued as a work of stimulus rather than of reference. P. M. DEANE.

8.—*Annual Epidemiological and Vital Statistics*, 1953. Geneva, World Health Organization, 1956. 571 pp. 11″. 50*s*.

This report is the sixth year book of epidemiology and vital statistics published by the World Health Organization. The series is a successor of the League of Nations publications and continues the epidemiological character of those reports. Death statistics are comparable between most countries regarding age and sex but for individual causes of death some reservations have to be made. Accuracy of diagnosis varies considerably between countries and in a number of countries

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

or sections of a country it is not necessary that the death be certified by a medical practitioner. At present the number of countries with really reliable health and vital statistics are in a minority. Despite the defects in the data useful comparisons can be made, especially between the more developed countries, and the report provides very useful material for all those interested in vital statistics. In the case of notifiable diseases the difficulties of making comparisons are very much increased and in many cases it is not justifiable to do more than compare the trends within the country. The difficulties arise from many sources and although the diseases are compulsorily notifiable the actual notifications often only represent a small fraction of the total cases occurring. Some diseases are only notifiable in a section of a country or only include a particular section of the population such as hospital in-patients and out-patients or persons attending hospitals or dispensaries.

No returns were made to the World Health Organization by Albania, Bulgaria, China (except Taiwan), Czechoslovakia, Hungary, Liberia, Poland, Roumania, Saudi Arabia, U.S.S.R. and Yemen. The report is in two sections.

Part I of the report deals with vital statistics and causes of death. The populations of the countries are given for ages under 1 year, 1–4 years and then in quinquennial age groups and a final group of 75 and over. The number of deaths is displayed in the same grouping except that the five year grouping is extended to age 84 and the final group is 85 and over. The general death rate per 1,000 of the population ranged in Africa from 5·9 among the European population of South West Africa to 22·3 in Egypt; in America from 6·3 in Alaska to 23·2 in Guatemala; in Asia from 6·3 in Israel to 13·6 in India; in Europe from 7·4 in Iceland to 13·0 in France and in Oceania from 5·6 in Hawaii to 10·8 among the Maoris of New Zealand.

The infant mortality per 1,000 live births varied in Africa from 25 among the European population of Southern Rhodesia to 179 in Egypt; in America from 28 in the United States to 139 in Barbados; in Asia from 34 in Formosa to 116 in India; in Europe from 17 in Iceland to 116 in Yugoslavia. The neo-natal rate is also given and in general where the infant mortality is low the deaths under 28 days of life account for more than half of the total infant deaths but where the infant mortality is high the neo-natal deaths form about one-third of the total infant deaths. For the chief infectious diseases of childhood, congenital malformations and diseases peculiar to early infancy the deaths have been tabulated in the following age groups: under 1 day, 1 to 6 days, 1 week to 1 month, 1 to 2 months, 3 to 6 months, 6 to 12 months, 1 to 2 years, 2 to 3 years, 3 to 4 years, 4 to 5 years.

The numbers of births and deaths are shown for each quarter of the years. The quarterly distribution of infant mortality and the number of neo-natal deaths are also given but this tabulation is not so complete, geographically, as that for births and deaths.

For 34 territories the number of deaths are tabulated by age and sex. The causes of death are classified under 50 headings. A special tabulation has been made of the deaths from tuberculosis and cancer which are given by age, sex and site. The number of deaths from tuberculosis is given for all forms, respiratory system and 10 other sites. Deaths from malignant neoplasms are given for 16 sites and all forms. For certain selected diseases, for about 20 countries, the death rates per 100,000 population have been tabulated by age and sex. The diseases selected were tuberculosis of respiratory system, malignant neoplasms, diabetes mellitus, diseases of the cardio-vascular system, vascular lesions affecting the central nervous system, chronic rheumatic heart disease, arteriosclerotic and degenerative heart disease, other diseases of the heart, hypertension with heart disease, hypertension without mention of heart disease, pneumonia, bronchitis, cirrhosis of the liver, nephritis and nephrosis, motor vehicle accidents, all other accidents, suicide and self-inflicted injury. The crude death rate from the various causes showed considerable variation between the different countries. The highest male rate for respiratory tuberculosis was 66·5 in Portugal and the lowest 8·5 in the Netherlands, the largest female rate was 46·7 in Japan and the lowest 4·7 in Australia, the death rate in England and Wales was 25·7 for males and 10·8 for females. The crude death rate from cancer was very high in England and Wales and only two territories, West Berlin and Austria, had a higher mortality. The death rate from the various forms of heart disease was relatively high in England and Wales. The most striking difference between the countries is the mortality from bronchitis. England and Wales had the highest crude death rates, 92·0 for males and 47·5 for females, and the next highest rate was that of Portugal, 25·0 for males and 17·2 for females. The lowest bronchitis death rate was that of the United States with a male rate of 2·5 and a female rate of 1·4. In about one half of the countries the death rate from bronchitis was quite small. France had a very high rate from cirrhosis of the liver, 35·2 for males and 20·7 for females, the rate for England and Wales was 3·1 for males and 2·2 for females. The highest male death rates from motor vehicle accidents were 36·6 in the United States and 35·2 in Australia and the lowest rate was 4·7 in Ceylon; the rate for England and Wales was 15·8. Austria had the highest male death rate, 66·8, from all other accidents and England

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and Wales had the lowest, 27·3. The highest suicide rates were those of West Berlin, 37·1 for males and 21·8 for females, and Switzerland, 34·4 for males and 10·0 for females. The suicide rate in England and Wales was low, 14·2 for males and 7·6 for females. Only 1·1 per cent. of the male deaths and 2·0 per cent. of the female deaths in England and Wales were due to senility, without mention of psychosis, and ill-defined and unknown causes, in France the percentages were 16·3 and 19·6, and in Portugal 16·8 and 22.9. In addition to these selected causes of death the crude death rate for each sex separately has been found for the remaining numerically important causes of death.

The second part of the report is concerned with the cases and deaths from notifiable diseases. For the 23 principal infectious diseases the cases and deaths are shown in either calendar months or four-week periods. For another 11 infectious diseases the annual figure is given. A greater sub-division of area has been made in the latter than in the former diseases. Tables for the pestilential communicable diseases do not include countries where no deaths or cases occurred. The incidence of diphtheria in England and Wales, 266 cases and 23 deaths, compare very favourably with the experience of other European countries; in Italy there were 15,749 cases with 857 deaths and Germany had 18,344 cases and 277 deaths. The worst European outbreaks of acute poliomyelitis were Italy with 5,010 cases and 375 deaths, Sweden with 5,090 cases and 216 deaths, and England and Wales with 4,547 cases and 320 deaths. W. J. MARTIN.

9.—*Books for All.* By R. E. Barker. Paris, UNESCO 1956. 102 pp. 10½″. 15*s*.

This survey of the world book trade was undertaken for Unesco by the Deputy Secretary of the Publishers Association of Great Britain to try to gauge the problems involved in the movement of books across the world. Needless to say, in a Unesco project, the aim is to promote freer circulation. Mr. Barker, in carrying out his task, has obviously been somewhat dismayed at the lack of facts to draw on. "A difficulty which has arisen in writing this study, however, is that statistics in the book trade, as in many other branches of commerce, are often inadequate."

The very definition of a book varies from country to country and, even then, is open to differing interpretations. There is also a fundamental difficulty in measuring book output: the accepted way in which statistics are collected is to record numbers of "titles" published without regard to the number of copies of each title printed, so that the limited edition of some erudite work ranks as one unit side by side with the best-seller. At the outset, then, just to get a measure of book output in different countries is by no means easy. Production in 60 countries, in terms of titles published, is set out in this survey and makes a useful record, slightly more extensive than that reproduced from Unesco annually in the United Nations Statistical Yearbook. Some useful figures are also given of the estimated average size of editions in different countries and, for twenty of them, book production is expressed in terms of total copies published instead of titles. This partial count is said to indicate that the total number of copies of books produced annually in the world approaches 5,000 million. However, this piece of deduction is by no means clear; perhaps some key piece of the jig-saw puzzle has been withheld, but the figuring given suggests a world total more in the region of 3,000 million. Deductions are not helped by the first two pictographs in the book which, excellent though they are in portrayal, purport to show on one page that ten countries are responsible for three-quarters of the world books (by titles), while on the facing page these same countries are shown to produce no more than two-thirds of the output of some 42 countries—again by titles.

Statistics about book production are weak: those about international trade in books are even more so. Mr. Barker has done what he can with them and managed to produce a broad picture but it cannot be accepted in too much detail. There are discrepancies between the statistics collected from the importing side and those collected from the exporting side which are not reconcilable in the usual manner. Also, it is not possible to gauge to what extent books are reaching the more remote corners of the world since these markets are usually merged in a group of "other countries" which may account for anything up to one-quarter of the trade.

Other statistics relate to the cost of carriage, the extent of translation into different languages, and summarized information about licensing systems and the imposition of import duties and sales taxes: (another Unesco publication has dealt with trade barriers in more detail). Suggestions are made for reducing obstacles to the flow of books but unfortunately nothing is said about the effect of censorship on this. There is much of interest in this book, but the overriding impression it will leave with any statistician is the need for steps to be taken to obtain better and standardized international statistics about books. MARJORIE DEANE.

10.—*Cahiers du Seminaire d'Econometrie No.* 3. *Les Modèles Économétriques.* Edited by René Roy. Paris, Centre Nationale de la Recherche Scientifique, 1955. 47 pp. 9½″. Fr. 750.

This is a useful collection of papers which were first read to the Econometric Seminar at the National Centre for Scientific Research in Paris.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Dr. Maillet presents two excellent survey articles, one on econometric models, the other on input-output theory. The latter includes some important new ideas on the relationship of Cobb-Douglas production functions to input-output theory, together with a detailed criticism of dynamic models. The former is supplemented by a lively discussion of the rival merits of recursive and interdependent models by Professor Wold and Dr. Malivaud, and of the use of static models in political economy by Professor Tinbergen.

The longest article in the collection is on the "Choice of Plant" by Dr. Janin of l'Electricité de France. His main thesis is that conventional methods of evaluating investment projects by their expected yield are impractical, and that simpler, *ad hoc*, methods based on the use of a few summary technical parameters and comparisons with a standard plant are to be preferred. A detailed analysis along these lines of choices made in the electricity industry would have been very interesting. Unfortunately Dr. Janin has preferred to state his argument in general terms, so that it sounds rather empty and disconnected.

Professor Marschak presents the only paper on economic theory in the collection. In this he develops various macromodels including money, some of which are shown to be internally inconsistent while others are not. Curiously enough there is no mention of Professor Patinkin's work.

W. M. GORMAN.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL NOTES

(1) British Official Statistics

The index of retail prices compiled by the Ministry of Labour and National Service which was 103 (January 17th, 1956 = 100) fell to 102 in June and remained at that figure in July and August. The fall in June was due mainly to reductions in the prices of potatoes, tomatoes and cabbage which were partly offset by increases in price for coal, eggs, bacon and butter. In July and August there were increases in prices for some commodities and reductions for others. The detailed figures of the weights used in calculating the index and the indices for the different commodity groups are as follows:

Date. Weights:	*Food* 350	*Alcoholic Drinks* 71	*Tobacco* 80	*Housing* 87	*Fuel and Light* 55	*Durable Household Goods* 66	*Clothing and Footwear* 106	*Transport and Vehicles* 68	*Miscellaneous goods* 59	*Services* 58	*All Items* 1,000
May 15th	104·2	100·7	105·3	102·7	98·0	101 2	100·4	101·9	101·6	102·1	102·5
June 12th	102·9	101·1	105·3	102·9	100·8	101·2	100·5	102·2	102·3	102·8	102·4
July 17th	101·1	101·9	105·3	103·1	100·9	101·3	100·5	102·4	103·1	105·1	102·0
Aug. 14th	101·7	101·9	105·3	103·2	101·1	101·3	100·6	102·4	103·2	105·2	102·3

A supplementary calculation shows that, on the basis of June 1947 = 100, the index at August 14th was 157, compared with 156 at July 17th, 157 at June 12th and May 15th. This supplementary calculation will not be made after the end of 1956.

The Ministry of Labour index of weekly wage rates which was 164 in May (June 30th, 1947 = 100) remained at that figure in June, July and August. A new calculation relating to manufacturing industries only shows a figure of 163 for the same dates. At the end of August the indices for all industries and services were 163 for men, 167 for women and 179 for juveniles. For manufacturing industries alone the figures were 161, 170 and 180.

The six-monthly enquiry into average earnings showed that in April, 1956, the average earnings per week in the industries covered by the enquiries, i.e. in the manufacturing and some of the principal non-manufacturing industries, were 235*s.* 4*d.* for men, 119*s.* 9*d.* for women; 100*s.* 6*d.* for youths and boys and 78*s.* 4*d.* for girls. These figures represent increases over October, 1938, of 241, 268, 285 and 323 per cent. For all workers combined the average was 197*s.* 9*d.*, an increase of 271 per cent. The index for April, 1956, compared with 100 at April 1947, was 191. This compares with the weekly wage *rate* index of 165. The average hours worked per week in April, 1956, were 46·7 (48·6 for men only). In manufacturing industries alone the average earnings in April, 1956, were 242*s.* 3*d.* for men and 120*s.* for women. The corresponding hourly earnings were 60·3*d.* and 34·7*d.*

The total working population and the numbers in civil employment in the four months ended July, 1956, were as follows:

(*Thousands*)

	Total Working Population			*Numbers in Civil Employment*		
Date	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*
April 1956 .	16,135	7,847	23,982	15,233	7,755	22,988
May 1956 .	16,138	7,846	23,984	15,254	7,763	23,017
June 1956 .	16,136	7,853	23,989	15,258	7,773	23,031
July 1956 .	16,130	7,869	23,999	15,253	7,789	23,042

The number of persons on the unemployment registers of the Employment Exchanges fell by 15,000 in June and rose by 16,000 in July and by 28,000 in August. The total for August represented 1·2 per cent of the estimated total of employees in Great Britain. The percentages in the different Regions ranged from 0·6 in the North Midlands to 2·1 in the Midlands and Wales and 2.2 in Scotland.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The following is a sex analysis of the figures:

Number of Unemployed Persons on the Registers of Employment Exchanges

Date	*Men and Boys*	*Women and Girls*	*Total*
May 14th, 1956 . .	153,388	84,043	237,431
June 11th, 1956 . .	146,206	76,371	222,577
July 16th, 1956 . .	161,020	77,117	238,137
Aug. 13th, 1956 . .	181,703	84,838	266,541

Of the total of 266,541 in August, 63,944 had been unemployed for not more than two weeks, 68,577 for more than two and less than eight weeks, and 86,368 for over eight weeks, while 47,652 were temporarily stopped.

The number of vacancies filled by the Employment Exchanges in the four weeks ended July 27th was 152,599 and there were 397,296 unfilled vacancies on the registers.

The numbers of insured persons absent from work owing to sickness, including self-employed as well as employed, were 881,000 in May, 823,000 in June, 793,200 in July and 789,200 in August. The numbers of employed persons absent owing to industrial injuries were 59,100, 56,400, 56,000 and 52,300.

A further Report to the Lord President of the Council and the Chancellor of the Exchequer by the standing Interdepartmental Committee on Social and Economic Research has been published (*Report of the Interdepartmental Committee on Social and Economic Research*, H.M.S.O. 2*s*. 6*d*. net.).

The main function of the Committee, whose members are drawn from the academic staffs of the Universities and from Government Departments, is to bring to the notice of Departments the potential value to research in the social sciences of the material which they collect, to suggest new methods and areas of collection, and to advise how the information gathered could be made available to research workers.

The third made by the Committee, this Report is primarily an account of work done during the last five years only. It does contain, however, a table which shows that at least 70 per cent. of the recommendations made to the Departments by the Committee during the whole period 1947–1955 were accepted.

As well as fostering direct collaboration between research workers and Government Departments on specific projects, the Committee has continued to assist research workers generally to find out what material is available in Government Departments. It has done so in two ways: by the addition of a booklet on Local Government Statistics to its series of Guides to Official Sources and by encouraging Departments to make available a limited number of copies of unpublished papers to University and other research libraries.

Recommendations made to Departments during the five years ranged widely. Patents, company records, insurance and bankruptcy statistics, as well as statistics relating to the administration of enemy property, are among the subjects on which recommendations were made to the Board of Trade; the Committee also conveyed its view to the Board on the question of publishing selected papers on foreign economic relations.

The Registrars General were advised on tables for information obtained by the 1951 Census on birthplace and nationality, education, classification by social class, the cross-classification of occupations within industries, fertility, housing and the characteristics of private households. The Committee also made suggestions for detailed census statistics of the large conurbations (Greater London, West Midlands, South East Lancashire, Merseyside, West Yorkshire, Tyneside and Central Clydeside).

The broad nature of the survey of Ministry of Food material did not preclude the Committee from making recommendations on points of detail. The Report refers to statistics of food supply and procurement, home production, processing and manufacture; stocks; the distribution and use of food; the administration of food controls; prices and margins; and household food surveys. In particular, the Committee point to the research interest in the continuance of the National Food Survey and advocate two publications on statistics of food and agriculture: a periodical for current figures and a further addition to the series of Guides to Official Sources.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(2) OTHER STATISTICS

The British Iron and Steel Federation Statistical Year Book for 1954 for Overseas Countries (451 pages, 15*s.*) has recently been published. It is a comprehensive summary of the world steel production figures. Though the detail for most of the countries is not carried beyond 1954, there are summary tables giving the world production of the various countries for iron ore, pig iron and crude steel up to the end of 1955. The value of the publication is greatly increased by the maps showing the location of the plants and raw material sources in each of the main producing countries.

The complementary volume for the United Kingdom, *Annual Statistics for 1955* (151 pages, 7*s.* 6*d.*), is also available. It is now published jointly by the Iron and Steel Board and the British Iron and Steel Federation following the new arrangements made for the collection and publication of United Kingdom statistics under the supervision of a Joint Committee of the Board and the Federation.

The Economic Commission for Europe (United Nations) has published a report in two volumes entitled *The Problem of Cost in the Inland Transport Industry*. The report is by a team of costing experts working under the auspices of the Inland Transport Committee. The work was carried out under the three headings of analysis, calculation, and use of costs. The main body of the report deals with problems of Railway Transport Cost, Road Transport Costs and Inland Water Transport Costs and then proceeds to deal with problems common to these three modes of transport.

There is a brief section upon statistical measurements in the transport industry and a further one upon railway statistics. The report (page 17) notes that "the transport statistics at present collected [are] far from complete in many respects". The second volume contains all the tabular and graphical annexes arranged in the same basic order as the sections of the first volume. Both volumes are substantial (about 250 pages of quarto paper) and there is a clear indication that others will follow in due course.

The United Nations *World Economic Survey* for 1955 (220 pp., 14*s.*) opens with a review of the last 10 years in compliance with a request of the Economic and Social Council for such a review in connection with the Council's 10th anniversary in 1955. It states that while the economic record of the past 10 years is superior to that of the first world war, especially in the expansion of world trade, the problem of mass poverty remains in a large part of the world as stubborn as ever and the gap in production between the developed and the underdeveloped areas of the world has been growing, and concludes that "significant as the post-war decade of growth has been . . . it cannot yet be said that the world has effectively come to grips with the problem of economic development". Part II of the Survey deals with the current situation, treating separately the private enterprise economies and the centrally planned economies. In the first group, expansion was the keynote of change in the developed countries while changes in the underdeveloped countries were uneven; in the second group, industrial expansion and better harvests contributed to higher activity. Planning continued to provide for rapid growth in heavy industry with greater emphasis on agriculture.

Two supplements are issued with the Survey. The first, *Economic Developments in the Middle East* 1954–1955 (155 pp., 11*s.*) gives a general description of trends in production, trade, finance and development programmes during 1954 and 1955 for 7 countries with a population of about 100 million. Owing to lack of data, little information is given for most of the chief oil-producing countries. The second, *Economic Development in Africa* 1954–1955 (100 pp., 7*s.*) reviews the growth of economic activity in this Continent (excluding Egypt, considered as part of the Middle East), noting differences in the economic structure of the principal regions; northern, tropical and southern Africa.

CURRENT NOTES

An International Conference on Operational Research is being organized by the Operational Research Society of the United Kingdom, the Operations Research Society of America, and the Institute of Management Sciences, U.S.A., and will be held at the University of Oxford from September 2nd to September 6th 1957. The object of the Conference will be to unify and extend

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the science of operational research. Papers will be presented in four sessions comprising: 1. The Common Themes in Operational Research; 2. Methodology; 3. and 4. Applications. The papers and proceedings of the conference will be published. Accommodation will be available in colleges. Visits are being arranged to establishments in England where operational research, electronic data processing, work study or other related work is being carried out.

Further information may be obtained from Mr. B. H. P. Rivett, Honorary Secretary, Operational Research Society of the United Kingdom, 2 Grosvenor Place, London, S.W.1.

The Trustees of the Houblon-Norman Fund invite applications for Fellowships or Grants in aid of Research into the working of industry and finance in Great Britain and elsewhere and the economic conditions affecting them.

Fellowships, which are intended for experienced research workers, are awarded for full-time research for one year, which may be renewed for a second year. Grants are awarded for either full-time or part-time work. The amount of the awards, which will take effect from the 1st October 1957, will depend upon the circumstances of the candidate and the probable expenses of his work.

Candidates must be British subjects normally resident in the United Kingdom.

Forms of application, which may be obtained from the Secretary of the Fund, c/o the Bank of England, London, E.C.2, should be returned not later than the 1st March, 1957.

The Trustees of the Houblon-Norman Fund, on the recommendation of their Advisory Committee, have made the following awards for 1956–57.

Research Grants.—Mr. J. D. Bailey, Graduate Student, University of Oxford, "British Investment in Australasian Financial Companies, 1870–1900"; Mrs. C. R. Fulop, Lecturer (Renewal), "Buying Practices of Retailers"; Mr. J. K. Horsefield, "Investigation into the connection between the Land-Bank Schemes of 1690–1720, the development of banking theory, and particularly the Recoinage of 1696"; Mr. J. Jewkes, Professor of Economic Organisation, University of Oxford, "Inventions and the Development of Inventions in the Aircraft Industry, 1900–1950"; Mr. R. V. Lennard, Historian (Renewal), "Rural England in the Feudal Age"; Miss C. A. P. Leubuscher, "Shipping Facilities available to the Overseas Trade of Colonial Territories"; Mr. A. L. Minkes, Tutor, Department of Extra-Mural Studies, University of Birmingham (Renewal), "The Economic Structure of the British Paint and Varnish Industry"; Miss J. E. Mitchell, Lecturer, University of Nottingham, "The Balance of Payments Crisis of 1951"; Mrs. C. S. Orwin, "English Farming, 1846–1913"; Mr. R. A. C. Parker, Lecturer, University of Manchester (Renewal), "18th and Early 19th Century Landed Estates"; Mr. H. W. Parris, W. E. A. Tutor, Northamptonshire (Renewal), "Board of Trade Railway Department and its influence on the Development of the Railway System in Great Britain"; Mr. A. J. Youngson, Fellow of Emmanuel College, Cambridge, "Comparative Study of Economic Development".

By a recent decision of the Council the prices of back volumes of the *Journal*, Series A and B, have been increased. The price per volume of available issues of Series A is now £3 up to and including Vol. 107 (1944) and £4 for Vol. 108 (1945) onwards. For Series B the price per volume is £3 for Vols. 1–7 (1934–1940/41) and £4 for Vol. 8 (1945) onwards. It is regretted that the concession hitherto given to Fellows to purchase back issues at a reduced rate can no longer be granted.

The Council offers the Frances Wood Memorial Prize of fifty guineas for competition in 1957. Competitors, who need not be Fellows of the Society, must be under the age of 35 on 1st October, 1957. The Prize is offered for the best investigation, on statistical lines, of any problem, to be chosen by the competitor, which bears directly or indirectly upon economic or social conditions.

Further particulars may be obtained from the Secretary of the Society.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

OBITUARY

JOHN WISHART, D.Sc., F.R.S.E.

It was a great shock to hear last July of John Wishart's death, due to a bathing accident at Acapulco in Mexico. Wishart was on leave of absence from Cambridge, and had been organizing the teaching of statistics in relation to agricultural research, a task undertaken on behalf of the United Nations Food and Agirculture Organization.

John Wishart was born in Perth on November 28, 1898, and educated at Perth Academy and Edinburgh University. His mathematical training led to an appointment as mathematics master at West Leeds High School from 1922 to 1924, but the first important step in his statistical career was his move to the Galton Laboratory, London, where he became a research assistant with Karl Pearson. The second step came in 1927 with an appointment in the statistical department at Rothamsted, with Dr. (now Sir) Ronald Fisher. Wishart's interest in and facility for computing, later indicated by his secretaryship of the British Association Mathematical Tables Committee from 1936 to 1947, and his thorough knowledge and experience of statistical techniques in agricultural research, must have owed much to these successive periods of apprenticeship in the Galton Laboratory and at Rothamsted, and to the consequent close association with Karl Pearson and with Fisher.

There is little doubt that Wishart's most active research years were spent at Rothamsted. He took his London D.Sc. in Statistics in 1928, and presumably included in his submitted publications his derivation of the generalized product moment distribution, published in *Biometrika* in the same year. The geometrical approach used by Fisher in his 1915 *Biometrika* paper on the sampling distribution of the correlation coefficient was again used by Wishart in this 1928 paper, the author fully acknowledging in his paper the help and inspiration Fisher had given him. Wishart's distribution was subsequently to play an important role in multivariate analysis, and it must have given Wishart considerable pleasure to see his name become permanently attached to the distribution he had first derived. Another theoretical line of research in which Fisher's influence can be traced, where Wishart showed particular and continued aptitude and interest, was in the sampling theory of estimated cumulants or semivariants. During his time at Rothamsted, Wishart also published applied or expository papers arising out of the institute's agricultural research, the joint papers with Fisher (1930) and Clapham (1929) [see selected Bibliography] being perhaps most representative.

In 1931 Wishart was appointed Reader in Statistics in the University of Cambridge, a post which he held until his death, apart from a period of war service with the Admiralty. G. U. Yule had just retired from full-time teaching, a little prematurely owing to ill-health, and after some negotiations the new Reader was set up with a small statistical laboratory in the School of Agriculture building, but also with the support and interest of the Mathematics Faculty. Wishart was keen from the beginning of his Cambridge appointment to introduce statistics to the mathematical undergraduates, and at once started a course of statistics among the lecture courses that could be offered for Schedule B of the Mathematical Tripos. This teaching was in addition to his commitments for the Faculty of Agriculture. The Economics Faculty, with which Yule had been associated, arranged separate courses for statistics for their own undergraduates.

At Cambridge Wishart had a long and at times rather frustrating struggle to get what he felt was sufficient recognition for statistics. It must be remembered that, apart from the Department of Statistics at University College, London, recognized university schools of so-called mathematical statistics hardly existed at this time. All the difficulties of knowing how to teach mathematicians statistics, or biologists mathematics, that have been so aired in recent years, Wishart found himself struggling with largely on his own. Many present university teachers (including the writer) owe much to the courses which he organized and taught with so much enthusiasm. They found him moreover a loyal friend, ready with his Scottish shrewdness to give them at the end of their student days sound advice on possible jobs; and they were always welcome visitors to the home he shared with his wife Olive and their two sons.

Wishart's dual responsibility to agriculture and mathematics was a delicate one which perhaps further contributed to his growing awareness that final stability at Cambridge had yet to be achieved; and his occasional impatience with those he felt were not wholeheartedly in sympathy with his proposals possibly contributed to a feeling of unease which his absence from Cambridge for national service during the war did not help to resolve. After the war, however, he found himself

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

able on the mathematical side to make further progress. This was in spite of the untimely deaths of Sir Arthur Eddington and Sir Ralph Fowler, who had always been sympathetic to statistics; the imperative war need for mathematicians or others with statistical qualifications no doubt helped Wishart with his post-war proposals. In the revision of the Cambridge Mathematical Tripos after the war, a course on "Random Variables" was instituted, as well as an expansion of the Schedule B (now called Part III) courses. A one-year Diploma course in Mathematical Statistics which Wishart initiated was available for the first time in 1947–8. Plans for a Statistical Laboratory within a Mathematical Institute were retarded through building difficulties, but temporary premises for such a Statistical Laboratory (coming entirely under the Faculty of Mathematics) were found in St. Andrews Hill, near the recently constituted Mathematical (Computing) Laboratory. In 1953 Wishart was appointed Director of the Statistical Laboratory, to which were also attached by this time three or four further members of the Mathematics Faculty concerned primarily with statistics.

Among other distinctions Wishart was a member of the International Institute of Statistics, a Fellow of the Institute of Mathematical Statistics, and an associate Editor of *Biometrika*. His services to the Society, of which he was a member of Council for many years, included an active part in the formation in 1933 of the Industrial and Agricultural Research Section, and he gave the second paper to the Section entitled "Statistics in agricultural research". In 1945 when Industrial Applications branched off from the Research Section, he was made the first chairman of the newly constituted Research Section.

We mourn this sudden loss of a friend and colleague who did so much for the development and teaching of statistics, especially mathematical statistics, in this country.

Selected Bibliography

(1928) "The generalized product moment distribution in samples from a normal multivariate population". *Biometrika*, **20**A, 32.
(1929) (with A. R. Clapham) "A study of sampling technique. The effect of artificial fertilizers on the yield of potatoes". *J. Agric. Sci.*, **19**, 600.
(1929) "A problem of combinatorial analysis giving the distribution of certain moment-statistics" *Proc. Lond. Math. Soc.*, **29**, 309.
(1930) (with R. A. Fisher), *The arrangement of field experiments and the statistical reduction of the results*. Imp. Bur. Soil Science, Tech. Comm. No. 10.
(1931) (with R. A. Fisher). "The derivation of the pattern formulae of two-way partitions from those of simpler patterns". *Proc. Lond. Math. Soc.*, Series 2, **33**, 195.
(1931) "The analysis of variance illustrated in its application to a complex agricultural experiment with sugar beet". *Arch. f. Pflanz.* **5**, 561.
(1934) "Statistics in agricultural research". *J. R. Statist. Soc.* Suppl. **1**, 26.
(1935) (with H. G. Sanders) *Principles and practice of field experimentation* (Empire Cotton-growing Corporation, London).
(1938) "Growth rate determinations in nutrition studies with the bacon-pig, and their analysis". *Biometrika*, **30**, 16.
(1939) "Statistical treatment of animal experiments". *J. R. Statist. Soc.* Suppl. **6**, 1.
(1939) "Some aspects of the teaching of statistics". *J. R. Statist. Soc.* **102**, 532.
(1940) *Field trials: their layout and statistical analysis. Imp. Bur. Plant Breeding and Genetics.*
(1947) "Statistical aspects of demobilization in the Royal Navy". *J. R. Statist. Soc.* **110**, 27.
(1948) "The teaching of statistics". *J. R. Statist. Soc.* **111**, 212.
(1950) *Field trials II: the analysis of covariance. Imp. Bur. Plant Breeding and Genetics.*

M. S. BARTLETT.

ARTHUR AVERY ADAMS

Arthur Avery Adams died suddenly on 3rd August, 1956, at the early age of 34, leaving a widow and two children, the second being born two days after his death. For eight years he had held a research post at the Department of Applied Economics, Cambridge, where he had acquired a unique reputation for original thought and inexhaustible patience and thoroughness in handling difficult and complex statistical tasks. For the last three years he had also been statistician to the London and Cambridge Economic Service, and his overhaul of the basic statistical tables in their Bulletin put users heavily in his debt.

His work as an economic statistician was largely on estimates of the national product in real

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

terms and related matters, on which he wrote many articles for the L.C.E. Bulletin. Immediately before his death he was in the middle of an article giving quarterly figures for the total supply and disposal of goods and services in the United Kingdom; fortunately (and typically) his papers were in sufficiently good order for this to be completed, and this pioneer work stands as the climax of his long fight against difficulties—including the paralysis which followed from an attack of poliomyelitis when he was 18, and which he bore with great cheerfulness and courage.

He was elected a Fellow of the Society in 1950 and only a few weeks before his death read a paper to the Cambridge Conference organized by the Study Section.

W. B. REDDAWAY.

MRS. J. R. N. STONE

Feodora Sonya Leontinoff, who died on June 29th, 1956, was born in 1909 in New Zealand. She obtained her education in this country and read philosophy at University College, where she was subsequently an assistant lecturer. In 1940 she was appointed Secretary to the National Institute of Economic and Social Research. She was elected a Fellow of the Society in 1941 and in the same year married Professor Richard Stone, to whom deep sympathy will be offered by all his friends in the Society.

Mrs. Stone had been for sixteen years the Secretary to the National Institute of Economic and Social Research.

She took up the position in 1940, when owing to the war the staff of the Institute had been greatly depleted and even its continued existence had become doubtful. She ensured its survival through the difficult war years and into the post-war period. With unflagging energy she strove with the problems of administration, money-raising, and recruitment and she had the satisfaction of seeing the Institute's reputation as a serious academic body, with a large and growing volume of work, firmly established.

She became a Fellow of the Royal Statistical Society in 1941 as representative of the Institute. In 1945 she was invited to serve on a Committee which the Society set up to consider the constitutional problem of the method by which its Council was elected. She took an active and constructive part in the deliberations of the Committee, acting as rapporteur on behalf of a subcommittee consisting of the non-Council members of the Committee. The recommendations of the Committee were adopted and form the basis of the present arrangements for the election of the Council.

Mrs. Stone devoted all her resources of intelligence, drive and executive ability, which were outstanding, to the advancement of social and economic knowledge in the United Kingdom; her premature death is a great loss to the Institute and to the wider world of academic learning which meant so much to her.

W. A. B. HOPKIN.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

STATISTICAL AND ECONOMIC ARTICLES IN RECENT PERIODICALS

UNITED KINGDOM—

Accounting Research—

July 1956—Credit costs and redemption practices in hire purchase: *J. Mars.* The social accounts of the United Kingdom for 1954 and 1955: *H. Burton.*

Aslib Proceedings—

August 1956—Statistical and economic information on textiles: *R. Robson.*

Heredity—

August 1956—ABO blood groups and racial characteristics in rural Wales: *I. M. Watkin.*

Operational Research Quarterly—

September 1956—Risk and hazard: *J. Cohen, E. J. Dearneley,* and *C. E. M. Hansel.* Statistics and hop growing experiments: *S. V. Cunliffe.* Intra-city trailer movement: *R. Vaswami.*

Review of Economic Studies—

Vol. XXIII, Part 3—The estimation of capital consumption in national accounting: *G. Stuvel.* The trader's behaviour in the market period: *S. Stykolt* and *A. Scott.* The Colonial Development Corporation (1948–54): *E. R. Wicker.* A note on the Colonial Development Corporation: *A. Hazelwood.* Comparative statics in linear programming and the Giffen paradox: *M. J. Beckmann.* A generalized comparative statics in linear programming: *M. J. Bailey.*

Cambridge Philosophical Society, Proceedings—

July 1956—Incomplete Bessel and Struve functions: *W. H. Steel* and *Joan Y. Ward.* The numerical evaluation of a class of integrals. II: *S. C. Das.* A new Monte Carlo technique: antithetic variates: *J. M. Hammersley* and *K. W. Morton.* General principles of antithetic variates: *J. M. Hammersley* and *J. G. Mauldon.* Exact tests for serial correlation in vector processes: *E. J. Hannan.* On a probability problem arising in the theory of counters: *L. Takács.* A measure of dispersion applied to cosmic-ray and other problems: *S. R. Broadbent.* The *G*-values of various games: *R. K. Guy* and *C. A. B. Smith.* Disjunctive games with the last player losing: *P. M. Grundy* and *C. A. B. Smith.*

Economic Journal—

September 1956—Competition in transport as an instrument of policy: *G. Walker.* Measurable utility and probability: a simplified rendering: *S. A. Ogza.* Wage policy and a labour standard: *J. H. Richardson.* Input-output analysis: an application: *A. A. Adams* and *I. G. Stewart.* Innovations and equilibrium growth: *H. J. Bruton.* The regional location of industry: *R. J. Nicholson.*

Incorporated Statistician—

January 1956—Economic accounting—a study in conceptional relationships: *F. Sewell Bray.* An approximation to the binomial: *J. Wishart.* Some aspects of classification data in market research: *T. Cauter.* Some uses of statistics in the brewing industry: *S. Cunliffe.*

Institute of Actuaries Students' Society Journal—

April 1956—The demographic situation in the United States of America and in England and Wales: *C. J. Thomas.*

Oxford University Institute of Statistics Bulletin—

May 1956—The variability of engineering earnings: *T. P. Hill* and *K. G. J. C. Knowles.* Small manufacturing businesses: *J. A. Bates* and *M. Stewart.* Capital intensity of industries: *V. V. Bhatt.* A "human needs" diet: spring 1956: *T. Schulz.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Population Studies—

July 1956—Fertility of polygamous marriages: *H. V. Muhsam.* A century of international mortality trends: II: *G. J. Stolnitz.* The trend of the birth rate in India, 1911–50: *A. Ghosh.* The birth statistics of Massachusetts during the nineteenth century: *R. Gutman.* Local population projections in England and Wales: *J. R. L. Schneider.*

INDIA—

Calcutta Statistical Association Bulletin—

June 1956—Optimum structure of rectangular sample units: *B. Ghosh.* Statistical tables and some recurrence in relations for *p*-statistics: *S. B. Chaudhury.* A model for perimeter errors: *B. Ghosh.* On a new class of partially balanced incomplete block designs: *H. R. Thompson.*

UNITED STATES—

American Economic Review—

June 1956—An essay on bargaining: *T. C. Schelling.* Population growth and the capital coefficient: *R. A. Gordon.* Union, non-union wage differentials: *J. E. Maher.* The inter-industry wage structure, 1899–1950: *D. E. Cullen.*

American Statistical Association, Journal—

June 1956—A sampling procedure for mailed questionnaires: *M. A. El-Badry.* The economic design of $\bar{X}$ charts used to maintain current control of a process: *A. J. Duncan.* Economically optimum acceptance tests: *J. V. Breakwell.* Analysis of sensitivity experiments when the levels of stimulus cannot be controlled: *A. Golub* and *F. E. Grubbs.* A note on uniformly best unbiased estimators for variance components: *F. A. Graybill* and *A. W. Wortham.* Some estimators in sampling with varying probabilities without replacements: *D. Raj.* The efficiencies of tests of randomness against normal regression: *A. Stuart.* A note on matrix inversion by the square root method: *D. Durand.* On approximating the point binomial: *M. S. Raff.* Fiscal-year reporting for corporate income tax: *W. L. Crum.*

September 1956—The optimum strategy in blackjack: *R. R. Baldwin, W. E. Cantey, H. Maisel* and *J. P. McDermott.* Two sequential tests against trend: *G. E. Noether.* Equivalence of two estimates of product variance: *D. W. Gaylor.* The economics of the President's economic reports: *A. C. Harberger.* Analytical graduation of fertility rates: *I. Lah.* Asymptotic efficiencies of a nonparametric life test for smaller percentiles of a gamma distribution: *J. E. Walsh.* Estimates of bounded relative error for the ratio of variances of normal distributions: *S. Reiter.* Machine computation of higher moments: *N. Jaspen.*

American Statistician—

June 1956—The role of statistics in shaping economic policy: *R. Bolling.* On the interpretation of the aggregate savings ratio: *R. Ferber.*

Annals of Mathematical Statistics—

June 1956—Alternative models for the analysis of variance: *H. Scheffé.* The theory of decision procedures for distributions with monotone likelihood ratio: *S. Karlin* and *H. Rubin.* On the power of certain tests for independence in bivariate populations: *H. S. Konijn.* The efficiency of some nonparametric competitors of the *t*-test: *J. L. Hodges, Jr.* and *E. L. Lehmann.* On regular best asymptotically normal estimates: *C. L. Chiang.* The large-sample power of rank order tests in the two-sample problem: *M. Dwass.* The estimation of the size of a stratified animal population: *D. G. Chapman* and *C. O. Junge, Jr.* On the design and comparison of certain dichotomous experiments: *R. N. Bradt* and *S. Karlin.* Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution: *D. Teichroew.* Estimation of location and scale parameters by order statistics from singly and doubly censored samples. Part I: *A. E. Sarhan* and *B. G. Greenwood.* A waiting line process of Markov type: *A. B. Clarke.* Accurate sequential tests on the mean of an exponential distribution: *G. E. Albert.* On the normal approximation to the hyper-geometric distribution: *W. L. Nicholson.* Charts of the power of the *F*-test: *M. Fox.* Estimating the parameters of a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

truncated gamma distribution: *D. G. Chapman.* Approximate upper percentage points for extreme values in multinomial sampling: *R. M. Kozelka.* An extension of the Kolmogorov distribution: *J. Blackman.* On the simultaneous analysis of variance test: *K. V. Ramachandran.* On the distribution of the likelihood ratio: *R. V. Hogg.* An application of Chung's lemma to the Kiefer-Wolfowitz stochastic approximation procedure: *C. Derman.* Some results on the restricted occupancy theory: *J. E. Freund* and *A. N. Pozner.* A vector form of the Wald-Wolfowitz-Hoeffding theorem: *D. A. S. Fraser.*

Biometrics—

June 1956—Confidence intervals for variance ratios specifying genetic heritability: *F. A. Graybill, F. Martin* and *G. Godfrey.* Programming analysis of variance for general purpose computors: *H. O. Hartley.* Note on fitting the multi-hit survival curve: *J. M. Gurian.* Quality evaluation by numerical and subjective methods with application to dried veneer: *W. G. Kauman, J. W. Gottstein* and *D. Lantican.* The estimation of age-specific infection rates from a curve of relative infection: *P. Whittle.* An evaluation of the removal method of estimating animal populations: *C. Zippin.* The concept of path coefficient and its impact on population genetics: *C. C. Li.* Estimation of the number of different classes in a population: *R. C. Lewontin* and *T. Prout.*

Industrial Quality Control—

July 1956—A check inspection and demerit rating plan: *H. F. Dodge* and *M. N. Torrey.* On the use of theory: *W. Edwards Deming.* Mail order acceptance sampling: *F. Streit.*

Journal of Political Economy—

June 1956—Observations on the economy of the USSR: *D. Gale Johnson.* Economic stabilization in a primary-producing country: *J. A. Hasson.* The baseball players' labor market: *S. Rottenberg.*

Management Science—

April 1956—General systems theory—a skeleton of science: *K. Boulding.* Present status of the work of managing: *H. F. Smiddy.* Management science in marketing: status and prospects: *M. Anschen.* A problem in optimal machine loading: *M. E. Salveson.* A computing procedure for a line balancing problem: *J. R. Jackson.* On the theory of dynamic programming—a warehousing problem: *R. Bellman.*

July 1956—Operations engineering: *E. R. Bowerman* and *S. B. Littauer.* Some models of inventory and an application: *E. Naddor.* A fruitful application of static marginal analysis: *H. W. Karr* and *M. A. Geisler.* A mathematical model for integrated business systems: *I. J. Lieberman.* An optimum geographical distribution of publicity expenditure in a private organization: *A. P. Zentler* and *D. Hyde.* Geographical distribution of production in multiple plant operations: *J. D. Dillon.*

Mathematical Tables and Other Aids to Computation—

April 1956—A modified Chebyshev-Everett interpolation formula: *G. A. Chisnall.* On a device for computing the $e_m(S_n)$ transformation: *P. Wynn.*

Operations Research—

June 1956—The development of operations research as a science: *R. L. Ackoff.* A study of sales operations: *A. A. Brown, F. T. Hulswit* and *J. D. Kettelle.* On staff organization: *R. H. Roy.* On operation analysis: *R. H. Roy.* The theory of search: part I, kinematic bases: *B. O. Koopman.*

August 1956—Education for operations research: *L. A. Brothers.* Fallacies in operations research: *B. O. Koopman.* The design of military supply tables for spare parts: *M. A. Geisler* and *H. W. Karr.* A two-phase method for the simplex tableau: *H. M. Wagner.* The relevance of costs in operations research: *M. W. Hoag.* The need for basic research in marketing: *J. B. McKitterick.* On the determination of optimum reserve generating capacity in an electric utility system: *E. L. Arnoff* and *J. C. Chambers.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Psychometrika—

June 1956—Optimal test length for maximum absolute prediction: *P. Horst* and *C. MacEwan.* A least squares solution for paired comparisons with incomplete data: *H. Gulliksen.* A non-parametric test of correlation using rank orders within subgroups: *W. S. Torgerson.* Maximizing test validity by item selection: *H. Webster.* The validity of the successive intervals method of psychometric scaling: *W. W. Rozeboom* and *L. V. Jones.* Relationships between two systems of factor analysis: *C. W. Harris.* Testing message diffusion in harmonic logistic curves: *S. C. Dodd.* A method for obtaining an ordered metric scale: *S. Siegel.*

Quarterly Journal of Economics—

August 1956—Income distribution, value judgments, and welfare: *F. M. Fisher.* Economic growth and income distribution: *W. P. Strassmann.* Profit inflation and industrial growth: the historic record and contemporary analogies: *D. Felix.*

Society of Actuaries, Transactions—

Eastern Spring Meeting Number, 1956—Frequency distribution of mortality costs: *J. M. Boermeester.*

AUSTRIA—

Statistische Vierteljahresschrift—

Vol. IX, Part 1—Die Risiktheorie in der Versicherungsmathematik: *L. Schmetterer.* Das Gesetz der grossen Zahl: *J. Pfanzagl.*

FRANCE—

Population—

April–June 1956—Un instrument d'étude des migrations intérieures: les migrations d'électeurs: *M. Croze.* Recherches relatives à la prévision de l'emploi: *J. Fourastié.* Evolution et perspectives régionales de population jusqu'en 1965: *R. Pressat.*

Revue de Statistique Appliquée—

Vol. IV, No. 2—La formation probabiliste et statistique de l'ingénieur: *G. Darmois.* L'interprétation statistique de la turbulence. État de la question: *A. Craya.* Le problème des corrélations dans la turbulence en soufflerie: *J. Bass.* Méthodes d'analyse de la structure fine des débits: *J. Larrieu.* L'utilisation des méthodes statistiques à l'Électricité de France pour la prévision des apports dans les réservoirs. Examen des résultats obtenus: *S. Ferry.* Des problèmes que pose le contrôle industriel: *J. Laroche* and *J. Mothes.*

GERMANY—

Allgemeines Statistisches Archiv—

Vol. 40, Part 2—Verteilungsfreie (nicht parametrische) Testverfahren in den Sozialwissenschaften: *O. Anderson.* Die repräsentative Bodenbenutzungserhebung in Baden Württemberg, 1956: *H. Wirth.* Zum Problem der Merkmalsauswahl für Trennverfahren (Barnard-Problem): *H. Baitsch.* Über Faustregeln zur optimalen Schichtung bei Normalverteilung: *H-J. Zindler.*

ITALY—

Giornale degli Economisti e Annali di Economia—

March–April 1956—Le leggi dello sviluppo pro capite nelle economie contemporanee: *G. Demaria.* Delusioni dell'Econometrica: *C. Gini.* Di una morfologia economica nella sua metodologia storiografica: *G. Palomba.*

Statistica—

April–June 1956—Sulle variabili divise in intervalli e le correzioni della media aritmetica e della varianza: *V. Castellano.* Come si risolve un problema di programmazione lineare mediante il metodo del potenziale logaritmico: *A. Predetti.* L'Istituto di Statistica della Facoltà di Economia e Commercio dell'Università di Firenze: *G. Parenti.* Intorno aï concetti di "popolazione giovane" e "popolazione vecchia": *R. Lenzi.* Aspetti attuali della congiuntura nel mercato cinematografico: *A. Bellettini.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

JAPAN—

Memoirs of the Faculty of Science, Kyusyu University, Series A (Mathematical)—

February 1956—Regular random functions and linear translatable stochastic functional equations: *T. Onoyama.*

Institute of Statistical Mathematics, Annals—

Vol. VII, No. 2—Decision rules, based on distance, for the problems of independence, invariance and two samples: *K. Matusita* and *H. Akaike.* Multidimensional central limit criterion in the case of bounded variances: *K. Takano.* Central convergence criterion in the multidimensional case: *K. Takano.* The effects of wide groupings on the distributions of array means and variances for correlated normal variables: *G. A. Baker.* Monte Carlo method applied to the solution of simultaneous linear equations: *H. Akaike.* An approximation method in numerical computation of the Leontief's open input-output model: *H. Kimura.* A generalization of Laplace criterion for decision problems: *H. Uzawa.* Shape and size distribution of carbon black when it is crushed: *A. Muta* and *I. Higuti.*

Vol. VII, No. 3—On the fundamental theorem for the decision rule based on distance: *K. Matusita* and *M. Motoo.* Some estimates which minimize the least upper bound of a probability together with the cost of observation: *H. S. Konijn.* Note on fitting a straight line when both variables are subject to error and some applications: *H. Hudimoto.* Some theorems on the sum of positive random variables: *M. Motoo.* On optimum character of von Neumann's Monte Carlo model: *H. Akaike.* On intertemporal efficiency conditions of capital accumulations (I): *H. Uzawa.* On the errors of outputs due to errors of technical coefficients in Leontief's open input-output models: *H. Kimura.* Determination of the target size by the indirect action of irradiation: *S. Sakino.*

NETHERLANDS—

Statistica—

Vol. 10, No. 2—Het probleem van de ruïnering der spellers: *M. Meinesz.* Betrouwbaarheidsintervallen voor de spreiding van een normale verdeling bepaald uit het gemiddelde van een aantal steekproefbreedtes: *G. Klerk-Grobben* and *H. D. Sandberg.* Significance of the smallest of a set of estimated normal variances: *R. Doornbos.*

SPAIN—

Trabajos de Estadistica—

Vol. VII, Part 1—Estudio de la evolución y relación de medidas antropométricas en los niños menores de un año: *S. Rios, J. Bejar, E. Jaso* and *A. Arbelo.* Tests no paramétricos en forma secuencial: *J. Romani.* La programmación lineal y la teoria económica. Aplicaciónes de la programación lineal: *J. Castañeda.*

SWEDEN—

Statistisk Tidskrift—

July 1956—Börsnoterade aktier och deras ägare: *F. Lublin.*

INTERNATIONAL—

International Labour Review—

September 1956—International movements of capital and labour since 1945: *Brinley Thomas.*

Revue de l'Institut International de Statistique—

Vol. 23, No. 1–3—La vingt-neuvième Session. Petropolis, 24 juin—2 juillet 1955: *G. Darmois.* The use of census of industrial production material for the estimation of productivity: *R. C. Geary* and *K. G. Forecast.* A sequential test with three possible decisions for comparing two unknown probabilities, based on groups of observations: *C. van Eeden.* Sur les théorèmes limites du calcul des probabilités: *D. Dugué.* The measurement of consumers' behaviour: *H. Theil.* La répartition des décès de cause " indéterminée ": *S. Ledermann.*

World Health Organization Bulletin—

Vol. 14, No. 5–6—Estimation of bacterial density of water samples: methods of attaining international comparability: *S. Swaroop.*

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

LIST OF ADDITIONS TO THE LIBRARY

Since the issue of Part III, 1956, the Society has received the publications enumerated below.

I.—OFFICIAL PUBLICATIONS

United Kingdom. *Board of Trade.* The report on the census of production for 1951. London, H.M.S.O., 1954–56. 12v. 157 pts. 11¼″. 259*s.* 6*d.*

Colonial Office. Colonial development and welfare, 1946–55. London, H.M.S.O., 1956. 40p. photos. illus. 10½″. 2*s.*

House of Commons. Library. A bibliography of parliamentary debates of Great Britain. London, H.M.S.O., 1956. 62p. diagr. 9¾″. 5*s.* (H.C. Library Document 2)

Lord President's Office. Report of the Interdepartmental Committee on Social and Economic Research. London, H.M.S.O., 1956. v, 33p. 9¾″. 2*s.* 6*d.* Third report.

Ministry of Agriculture, Fisheries and Food. Agricultural Marketing Acts, 1931–1949. Proposed British egg marketing scheme: report by G. G. Baker . . . on the public inquiry into objections (18th June, 1956). London, H.M.S.O., 1956. 172p. 9¾″. 6*s.* (Cmd. 9805)

Ministry of Health. National Health Service Act, 1946. Hospitals' directory, England and Wales, 1955. London, H.M.S.O., 1956. vi, 287 p. 10″. 15*s.*

Ministry of Housing and Local Government. Local government: areas and status of local authorities in England and Wales. London, H.M.S.O., 1956. 17p. 9¾″. 1*s.* (Cmd. 9831)

Ministry of Labour and National Service. Port transport industry: report of a committee appointed on 27th July, 1955, to inquire into the operation of the Dock Workers (Regulation of Employment) Scheme, 1947. London, H.M.S.O., 1956. iii, 63p. 9¾″. 3*s.* 6*d.* (Cmd. 9813)

Monopolies and Restrictive Practices Commission. Report on the supply of certain rubber footwear. London, H.M.S.O., 1956. vi, 107p. 2 fold. tables. 9¾″. 5*s.* (H.C. 328)

Report on the supply of hard fibre cordage. London, H.M.S.O., 1956. v, 143p. 9¾″. 4*s.* 6*d.* (H.C. 294)

Northern Ireland. General Register Office. The Ulster year book: the official year book of Northern Ireland, 1956. Belfast, H.M.S.O., 1956. xlv, 345p. ports. plates. diagrs. fold. map. 9¾″. 5*s.*

Royal Commission on Common Land. Minutes of evidence. 1–4. London, H.M.S.O., 1956. 4 pts. 9¾″. 10*s.*

Scotland. Department of Agriculture. Report of the Hill Lands (North of Scotland) Commission. Edinburgh, H.M.S.O., 1956. 22p. 9¾″. 1*s.* 3*d.* (Cmd. 9759)

Scotland. General Registry Office. Census 1951. Scotland. Vol. V. Fertility of marriage. Edinburgh, H.M.S.O., 1956. xvi, 86p. 13″. 25*s.*

Scottish Milk Marketing Board. Agricultural Marketing Acts, 1931 to 1939. Scottish milk marketing scheme, 1933 (incorporating amendments made in the scheme up to and including 1st May, 1956). Edinburgh, H.M.S.O., 1956. 28p. 9¾″. 1*s.* 9*d.*

Scottish Office. Scottish administration: a handbook. Revised ed. Edinburgh, H.M.S.O., 1956. 59p. 9¾″. 2*s.* 6*d.*

Scottish Office & Ministry of Agriculture, Fisheries and Food. Report of the Reorganisation Commission for pigs and bacon. London, H.M.S.O., 1956. viii, 101, [6] p. 2 fold. maps. 9¾″. 5*s.* 6*d.* (Cmd. 9795)

Australia. *Bureau of Agricultural Economics.* The Australian sheep industry survey, 1954. Pastoral zone. iv, 60p. Tasmania. [vi], 25p. Canberra, 1956. 9¾″.

1953 dairy survey. Costs, incomes and management problems on dairy farms in Victoria. Canberra, 1956. [x], 120p. 9½″.

Brazil. *Conselho Nacional de Estatistica. Serviço Grafico.* Bibliografia geografico-estatistica brasileira. Vol. I. 1936–1950. Rio de Janeiro, 1956. xlv, 362p. [6] l. 9¾″.

Canada. *Dominion Bureau of Statistics. Industry and Merchandising Division.* Energy sources in Canada: commodity accounts for 1948 and 1952. Ottawa, 1956. 59p. 11″. $1. (Ref. Pap. 69)

Customs Co-operation Council. Nomenclature for the classification of goods in customs tariffs, 1955. [ii], viii, 197p. 30*s.* French and English. Alphabetical index to the goods specified in the Brussels nomenclature, 1955. [iii], 150p. 30*s.* Explanatory notes to the Brussels nomenclature, 1955. 3v. 147*s.* Brussels, 1955. 11½″.

Finland. *Statistiska Centralbyrån.* Väestötilastoa . . . 1950 population census. Vol. I. Population, number, age and language. 69p. Vol. IV. Population by industry and industrial status: detailed classification. 186p. Helsinki, 1956. 11¾″. (Official Statistics of Finland VI. C 102)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Frankfurt am Main. *Statistisches Amt.* Die Landtagswahl am 28. November, 1954 in Frankfurt am Main: neue Beiträge zu einer statistischen Analyse des Wahlgeschehens; von Professor Dr. Rudolf Gunzert. Frankfurt, 1956. vi, 110p. 11¾″. (Sonderheft 5, *Statist. Monatsber., Frankfurt am Main*, 1956, **18**)

International Labour Office. Social aspects of European economic co-operation: report by a group of experts. Geneva, 1956. vii, 179p. 9½″. 9*s*. (Studies and Reports, N.S. 46)

Italy. *Istituto Centrale di Statistica.* L'attività dell'Istituto Centrale di Statistica nell'anno 1955: relazione al Consiglio superiore di statistica del Prof. Lanfranco Maroi, Presidente dell' Istituto. Roma, 1956. 38p. 10¼″.

IX censimento generale della popolazione e rilevazione delle abitazioni. III censimento generale dell'industria e del commercio, 4 e 5 novembre, 1951. Primi risultati generali dei censimenti: popolazione residente e presente—abitazioni e stanze—ditte, unità locali e addette. Dati provvisori. Roma, [n.d.] 188p. 10¼″. *L*500.

Elezione della camera dei deputati, 7 giugno, 1953. Vol. II. Voti di lista e voti di preferenza—candidati ed eletti. Roma, 1956. 143p. 10½″. *L*1000.

Japan. *Bureau of Statistics.* Population census of 1950. Population map. Part I. 1. Population density by *Shi*, *Ku*, *Machi*, and *Mura*. 2. Population dot map. 3. Population isopleth map. Tokyo, Bureau of Statistics, 1956. 5 sheets. 43″ × 31¼″.

Linz. *Statistisches Amt.* Beiträge zur Raumforschung im Grossraum Linz, 1955. Linz, 1956. 98p. fold. maps & tables. 8¼″. Preprinted from *Statistisches Jahrbuch der Stadt Linz*, 1955.

New Zealand. *Department of Statistics.* Population census, 1951. Vol. VIII. General report. Wellington, 1956. 195p. 12¼″.

Norway. *Statistisk Sentralbyrå.* Avskrivning og skattlegging. Depreciation and taxation. Oslo, 1956. 85p. 9½″. (Samfunnsøkon. Stud. 5). English summary.

Organisation for European Economic Co-operation. Maritime transport. Paris, 1956. 75p. 9½″. 9*s*. (Trends in Economic Sectors)

O.E.E.C. bibliographies. No. 2. International trade: theory of international trade, international trade policy, trade liberalisation, 1950–1955. Paris, [1956]. 85p. 9½″. 5*s*.

Sweden. *Statistiska Centralbyrån.* Folkräkningen den 31 december, 1950. VIII. Urvalsundersökningar: statistiken över förmögenhet ... Sample surveys: statistics of property. Stockholm, 1956. xiv, 66°, 77p. 9½″. (Sveriges Officiella Statistik)

United Nations. Proceedings of the World Population Conference, 1954, Rome, 31 August–10 September, 1954. Papers: Vol. 1 (Meetings 2, 4, 6 and 8). New York, 1956. London, H.M.S.O. xiii, 1040p. 11¾″. 11*s*. Various languages, summaries in English and French.

Department of Economic and Social Affairs. Economic developments in Africa, 1954–1955: supplement to *World Economic Survey, 1955*. New York, 1956. vii, 100p. 11″. 7*s*.

Economic developments in the Middle East, 1954–1955: supplement to *World Economic Survey, 1955*. New York, 1956. viii, 151p. 10¾″. 11*s*.

United States. *Bureau of the Census.* Bureau of the Census manual of tabular presentation: an outline of theory and practice in the presentaton of statistical data in tables for publication; prepared by Bruce L. Jenkinson, Chief, Statistical Abstract Unit. Washington, Superintendent of Documents, 1949, xiv, 266p. 9¼″. $1.75.

National Bureau of Standards. Tables of coefficients for the numerical calculation of Laplace transforms; by Herbert E. Salzer. Washington, Superintendent of Documents, 1953. ii, 36p. 10¼″. 25*c*. (Appl. Math. Ser. 30).

National Science Foundation. Research by cooperative organizations: a survey of scientific research by trade associations, professional and technical societies, and other cooperative groups, 1953; prepared for the National Science Foundation by Battelle Memorial Institute. Washington, Superintendent of Documents, 1956. viii, 47p. 10¼″. 35*c*. (Nat. Sci. Stud. NSF 56–12)

West Bengal. *State Statistical Bureau.* Economic survey of small industries, 1954. State of West Bengal. Alipore, Superintendent, Government Printing, 1956. [v], 205p. 9¾″. 9*s*. 9*d*.

II.—AUTHORS AND MISCELLANEOUS

Abrams, I. J. A note on the optimal character of the (*s*, *S*) policy in the inventory problem. Berkeley & Los Angeles, University of California Press, 1956. London, Cambridge University Press. [ii], 185–194p. 10¼″. 50*c*. (Univ. Calif. Publications in Statistics, Vol. 2, No. 9)

Advertising Association. A practical guide to the promotion of exports: a special issue of the journal of the Advertising Association; guest editor Roger Falk, O.B.E. London, 1956. [i], 31p. 8½″. gratis.

Association of Certified and Corporate Accountants. Library catalogue. London, The Association, 1956. [iii], 106p. 9¼″. 1*s*. 6*d*.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Bailey, N. T. J. On estimating the latent and infectious periods of measles. I. Families with two susceptibles only. *Biometrika* (1956), **43**, 15–22. 10¾″.

Science, statistics and operational research. *Research* (1956), **9**, 202–207. 9¾″.

Barna, Tibor, *editor.* The structural interdependence of the economy: proceedings of an international conference on input-output analysis, Varenna, 27 June–10 July, 1954. New York, Wiley, [1956]. Milano, Giuffrè. ix, 429p. 9½″.

Barnard, Mildred M. & Ditchburne, Nell. Elementary statistics for use in timber research. Melbourne, Commonwealth Scientific and Industrial Research Organization, 1956. 32p. 9¼″.

Bartlett, M. S. Deterministic and stochastic models for recurrent epidemics. *Proc. 3rd Berkeley Symp. Math. Statist., Dec. 1954, June–July 1955*, **4**, 81–109. 10″.

Bellamy, Dennis. The domestic consumer. Torquay, British Electrical Power Convention, 1956. [ii], 77p. 8½″. (8th Annual Convention, Torquay, May 28th–June 1st, 1956)

Bennett, B. M. The cumulants of a sample mean from a finite population of first *n* integers. *Trab. Estadist.* (1955), **6**, 31–32. 9½″.

Estimation of LD50 by moving averages. *J. Hyg., Camb.* (1952), **50**, 157–164. 10″.

Estimation of means on the basis of preliminary tests of significance. *Ann. Inst. Statist. Math.* (1952), **4**, 31–43. 10¼″.

Note on a solution of the generalized Behrens-Fisher problem. *Ann. Inst. Statist. Math.* (1951), **2**, 87–90. 10¼″.

Note on the moments of the logarithmic non-central χ^2 and z distributions. *Ann. Inst. Statist. Math.* (1955), **7**, 57–61. 10¼″.

On the cumulants of the logarithmic generalized variance and variance ratio. Skand. Aktu-Tidskr. (1955), 17–21. 9¼″.

On the joint distribution of the mean and standard deviation. *Ann. Inst. Statist. Math.* (1955), **7**, 63–66. 10¼″.

Some further extensions of Fieller's theorem. *Ann. Inst. Statist. Math.* (1953), **5**, 103–106. 10½″.

Blanco, Eugenio A. La politica presupuestaria, la deuda publica y la economia nacional: conferencia dictada en la Escuela Superior de Guerra, el 27 de abril de 1956. Buenos Aires, Ministerio de Hacienda, 1956. 55p. 10½″.

Brandt, Karl. The orientation of agricultural economics. *J. Farm. Econ.* (1955), **37**, 793–806. 9″.

Brearley, A. & Cox, D. R. An outline of statistical methods for use in the textile industry. 3rd ed. Leeds, Wool Industries Research Association, [1954]. 77p. 9¾″.

Cambridge. University. *Department of Applied Economics.* On criteria for descriptions of income distribution; by J. Aitchison and J. A. C. Brown. *Metroeconomica* (1954), **6**, 88–107. 9½″. (Repr. Ser. 108)

The implications of early national income estimates for the measurement of long-term economic growth in the United Kingdom; by Phyllis Deane. *Econ. Devpt. & Cult. Change* (1955), **4**, 3–38. 9½″. (Repr. Ser. 109)

The grouping of observations in regression analysis; by S. J. Prais and J. Aitchison. *Rev. Int. Statist. Inst.* (1954), **22**, 1–22. 10½″. (Repr. Ser. 110)

Some difficulties in a certain concept of community indifference; by R. W. Pfouts. *Metroeconomica* (1955), **7**, 16–26. 9½″. (Repr. Ser. 111)

National income and related statistics; by J. E. G. Utting. *J. Roy. Statist. Soc. A* (1955), **118**, 434–451. 9¾″. (Repr. Ser. 112)

On the distribution of a positive random variable having a discrete probability mass at the origin; by John Aitchison. *J. Amer. Statist. Ass.* (1955), **50**, 901–908. 9¼″. (Repr. Ser. 113)

Cheeseman, E. A., Crozier, C. M. E. & Merrett, J. D. Data on linkage in man: hereditary deafness in Northern Ireland. *Ann. Hum. Genet.* (1956), **21**, 59–66. 10¾″.

Cherry, Colin, *editor.* Information theory: papers read at a symposium on "information theory" held at the Royal Institution, London, September 12th to 16th, 1955. London, Butterworth, 1956. New York, Academic Press. xii, 401p. illus. 9¾″. 70*s.*

Croxton, Frederick E. & Cowden, Dudley, J. Applied general statistics. 2nd ed. New York, Prentice-Hall, 1955. London, Pitman. xvi, 843p. illus. 9″. 55*s.*

Dandekar, V. M. Use of food surpluses for economic development. Poona, Gokhale Institute of Politics and Economics, 1956. xiv, 153p. 9¾″. 7*s.* 6*d.* (Publ. 33)

Davies, J. O. F. & Barr, A. Sickness and non-sickness absence among nurses. *Mon. Bull. Minist. Hlth. Lab. Serv.* (1956), **15**, 158–194. 9¾″.

Davies, Joseph S. The economic potentials of the United States. *Columb. Univ. Bicentenn. Conf. III.* (1954). 5 l. 11″.

Edwards, D. A. & Moyal, J. E. Stochastic differential equations. *Proc. Camb. Phil. Soc.* (1955), **51**, 663–677. 10¼″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Farnsworth, Helen C. International wheat agreements and problems, 1949–56. *Quart. J. Econ.* (1956), **70**, 217–248. 9″.

Farnsworth, Helen C. & Jones, William O. Response of wheat growers to price changes: appropriate or perverse? *Econ. J.* (1956), **66**, 271–287. 17p. 9¾″.

Fegiz, P. Luzzatto. Il volto sconosciuto dell'Italia: dieci anni di sondaggi Doxa. Milano, Giuffrè, 1956. xxvii, 1353p. 9½″. *L*6000.

Finney, D. J. Cross-over designs in bioassay. *Proc. Roy. Soc. B* (1956), **145**, 42–61. 10″.

Multivariate analysis and agricultural experiments. *Biometrics* (1956), **12**, 67–71. 9¼″.

Finney, D. J. & Outhwaite, Anne D. Serially balanced sequences in bioassay. *Proc. Roy. Soc. B* (1956), **145**, 493–507. 10″.

Fisher, *Sir* Ronald A. Statistical methods and scientific inference. Edinburgh & London, Oliver & Boyd, 1956. viii, 175p. 8¾″. 16*s*.

Fix, Evelyn. Tables of noncentral χ^2. Berkeley & Los Angeles, University of California Press, 1949. London, Cambridge University Press. [ii], 15–19p. 10¼″. 25*c*. (Univ. Calif. Publications in Statistics, Vol. 1, No. 2)

Good, I. J. A classification of rules for writing informative English. *Methodos* (1955), **7**, 193–200. 9½″.

Good, I. J. & Toulmin, G. H. The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika* (1956), **43**, 45–63. 10¾″.

Hall, H. Duncan & Wrigley, C. C. Studies of overseas supply; by H. Duncan Hall and C. C. Wrigley, with a chapter by J. D. Scott. London, H.M.S.O. & Longmans, Green, 1956. xi, 537p. fold. table. 9½″. 37*s*. 6*d*. (History of the Second World War, U.K. Civil Series)

Hammersley, J. M. The area enclosed by Pólya's walk. *Proc. Camb. Phil. Soc.* (1956), **52**, 78–87. 10¼″.

Conditional Monte Carlo. *J. Ass. Computing Mach.* (1956), **3**, 73–76. 10″.

Hammersley, J. M. & Mauldon, J. G. General principles of antithetic variates. *Proc. Camb. Phil. Soc.* (1956), **52**, 476–481. 10¼″.

Hammersley, J. M. & Morton, K. W. A new Monte Carlo technique: antithetic variates. *Proc. Camb. Phil. Soc.* (1956), **52**, 449–475. 10¼″.

Hammersley, J. M. & Nelder, J. A. Sampling from an isotropic Gaussian process. *Proc. Camb. Phil. Soc.* (1955), **51**, 652–662. 10¼″.

Hammond, R. J. Food. Vol. II. Studies in administration and control. London, H.M.S.O. & Longmans, Green, 1956. xiii, 835p. 9½″. 50*s*. (History of the Second World War, U.K. Civil Series)

Food and agriculture in Britain, 1939–45: aspects of wartime control. Stanford, California, University Press, 1954. xii, 246p. 9″. $5. (Stanford University Food Research Institute Publication)

Herdan, G. Language as choice and chance. Groningen, Noordhoff, 1956. xiii, 356p. 9¾″. $8.

Hill, A. Bradford. Statistics in medical research. *Triangle* (1956), **11**, 208–215. illus. 9¾″.

Hodges, J. L., *jr*. The choice of inspection stringency in acceptance sampling by attributes. Berkeley & Los Angeles, University of California Press, 1949. London, Cambridge University Press. [ii], 14p. 10¼″. 25*c*. (Univ. Calif. Publications in Statistics, Vol. 1, No. 1)

Hoel, Paul G. Answers to exercises in *Introduction to mathematical statistics*. 2nd ed. New York, Wiley, [n.d.] London, Chapman & Hall. 11p. 8¾″. 5*s*.

Hulton Research. The Hulton readership survey, 1956. Condensed report; produced by British Market Research Bureau Limited for Hulton Research. London, Hulton Press, 1956. [i] 36 [3]p. 8″ × 3½″.

Ilersic, A. R. Statistics and their application to commerce. Revised 11th ed. London, H.F.L. (Publishers), 1956. xii, 401p. 8½″. 25*s*.

Italy. *Cassa per il Mezzogiorno*. La Cassa per il Mezzogiorno: primo quinquennio, 1950–1955. Roma, [printed] 1955. xii, 583p. photos, maps and summaries in English, French, German and Spanish, on unnumbered pages. 12½″.

Johnson, Robert E. & Morris, Doris N. Guide to elementary statistical formulas. New York & London, McGraw-Hill, 1956. [xii], 102p. 9″. 22*s*. 6*d*.

Jones, William O. Farm price policy and the surpluses. *Pacific Spectator* (1956), **10**, No. 3. 12p. 9″.

Kemp, C. D. & Kemp, A. W. The analysis of point quadrat data. *Aust. J. Bot.* (1956), **4**, 167–174. 9½″.

Kjaer-Hansen, Max. Selling costs and advertising expenditure in Scandinavia: an analysis and evaluation of the costs of commercial advertising in Denmark, Finland, Norway and Sweden. Copenhagen, Federation of Nordic Marketing Associations, 1956. 40p. 9¾″. (Publ. 1)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Klein, Josephine. The study of groups. London, Routledge & Kegan Paul, 1956. ix, 200p. 8½″. 21*s*.

Leander, Erik K. & Finney, D. J. An extension of the use of the χ^2 test. *Appl. Statist.* (1956), **5**, 132–136. 10″.

LeCam, Lucien. On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. Berkeley & Los Angeles, University of California Press, 1953. London, Cambridge University Press. [ii], 277–329p. 10¼″. 75*c*. (Univ. Calif. Publications in Statistics, Vol. 1, No. 11)

Man-made Fibres Producers' Committee. Man-made fibres. [Manchester, 1956.] iv, 26 l., 7 fold.diagrs. 9½″.

Martin, W. J. The distribution in England and Wales of mortality from coronary disease. *Brit. Med. J.* (1956), **1**, p.1523. 8p. 8½″.

A study of sex, age, and regional differences in the advantage of rural over urban mortality. *Brit. J. Prev. Soc. Med.* (1956), **10**, 88–91. 9¾″.

Matsukawa, Shichiro. Origin and significance of political arithmetic. *Ann. Hitotsubashi Acad.* (1955), **6**, 53–79. 10″.

Meyer, Herbert A., *editor.* Symposium on Monte Carlo methods: held at the University of Florida ... March 16 and 17, 1954. New York, Wiley, 1956. London, Chapman & Hall. xvi, 382p. 11″. 60*s*.

Michalup, Erich. Algunos aspectos demograficos de la poblacion de Venezuela. *Memorias 29a Sesion del Inst. int. Estadist.*, 1955. 7p. 10¾″.

Metodos de enseñanza estadistica y planes de estudio en la Universidad Central de Venezuela. Brasil, III Conferencia Interamericana de Estadistica, 1955. 5 l. 11¼″. (Documento de referencia 3387 Esp.)

Mogey, J. M. Family and neighbourhood: two studies in Oxford. London, Oxford University Press, 1956. xvi, 181p. 2 plates. 8½″. 30*s*.

Moroney, M. J. Facts from figures. 3rd ed. Harmondsworth, Penguin Books, 1956. viii, 472p. 7¼″. 5*s*.

Muranjan, S. K. Reflections on economic growth and progress. Poona, Gokhale Institute of Politics and Economics, 1956. 24p. 9¾″. 1*s*. 6*d*. (R. R. Kale Memorial Lecture)

National Bureau of Economic Research. The pattern of financial asset ownership: Wisconsin individuals, 1949; by Thomas R. Atkinson. Princeton, University Press, 1956. London, Oxford University Press. xviii, 176p. 9″. 30*s*. (Financial Research Program)

National Institute of Economic and Social Research. Some aspects of post-war company finance: an analysis of tabulations of company accounts published in *The Economist*; by Andrei Luboff. *Accounting Res.* (1956), **7**, 154–200. 46p. 9¾″. (Repr. Ser. 7)

Pearce, Henry George. Value, normal and morbid: an exposition of the Georgian theory of economic value. Sydney, The Standard, 1946. x, 272p. 7″.

Pol, B. van der & Bremmer, H. Operational calculus based on the two-sided Laplace integral. 2nd ed. Cambridge, University Press, 1955. xiii, 413p. 9¾″. 60*s*.

Political and Economic Planning. The British film industry: a report on its history and present organisation, with special reference to the economic problems of British feature film production. London, P.E.P., 1952. 307p. 9″.

Priestley, M. B. The analysis of stationary time series. II. Estimation of power spectra. London, Ministry of Supply, 1956. 18 l. 13″. (Tech. Note M.S.29)

Proceedings of symposia in applied mathematics. Vol. VI. Numerical analysis. New York & London, McGraw-Hill, for the American Mathematical Society, 1956. vi, 303p. 10″. 73*s*.

Reading. University. *Department of Agricultural Economics.* The milk-selling farms of Buckinghamshire; by R. S. Cook and A. H. Gill. Reading, 1956. [vii], 80p. 9¾″. 8*s*. 6*d*. (Misc. Stud. 9)

Scheffé, Henry. A "mixed model" for the analysis of variance. *Ann. Math. Statist.* (1956), **27**, 23–36. 10″.

Snodgrass, Burns. Teach yourself the slide rule. London, English Universities Press, 1955. 207p. 7″. 6*s*.

Spicer, C. C. The estimation of the parameters of an exponentially declining population. *J. Hyg., Camb.* (1956), **54**, 304–310. 10″.

Stolnitz, George J. Life tables from limited data: a demographic approach. Princeton, University, Office of Population Research, 1956. xii, 164p. 9″. $4.

Turnbull, H. W. The theory of determinants, matrices and invariants. 2nd ed. London, Blackie, 1950. xvi, 368p. 8½″. 35*s*.

Wishart, John. χ^2 probabilities for large numbers of degrees of freedom. *Biometrika* (1956), **43**, 92–95. 10¾″.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

REVENUE OF THE UNITED KINGDOM

Net Produce in Quarters of 1955, *and the Financial Years ended March* 31, 1955/56, 1954/55, 1953/54, 1952/53

£ thousand

	Quarters ended				Total for Calendar Year 1955
	March 31, 1955	June 30, 1955	Sept. 30, 1955	Dec. 31, 1955	
Ordinary Revenue—					
Income tax	1,217,471	239,516	238,350	242,282	1,937,619
Surtax	84,600	19,400	9,800	19,000	132,800
Death duties	50,900	46,100	45,500	40,500	183,000
Stamps	19,530	16,500	20,900	16,400	73,330
Profits tax and excess profits tax	42,000	33,100	59,500	58,700	193,300
Excess profits levy	14,450	9,000	4,100	2,900	30,450
Other Inland Revenue duties	450	150	100	100	800
Total Inland Revenue	1,429,401	363,766	378,250	379,882	2,551,299
Customs	257,362	282,177	286,029	309,927	1,135,495
Excise	200,569	176,920	220,560	231,720	829,769
Total Customs and Excise	457,931	459,097	506,589	541,647	1,965,264
Motor vehicle duties	58,381	9,362	7,961	6,855	82,559
Post Office (net receipt)	3,900	200	—	3,400	7,500
Broadcast receiving licences	7,300	4,300	4,800	8,700	25,100
Receipts from sundry loans	2,263	2,837	19,776	4,829	29,705
Miscellaneous receipts	63,135	18,556	41,614	83,243	206,548
Total Ordinary Revenue	2,022,311	858,118	958,990	1,028,556	4,867,975
Self-balancing Revenue —					
Post Office	59,000	62,800	63,000	66,400	251,200
Income tax deducted from excess profits tax, post-war refunds	2,163	795	948	225	4,131

£ thousand

	Year 1955/56	Year 1954/55	1955/56 (compared with 1954/55)		Corresponding Years	
			Increase	Decrease	1953/54	1952/53
Ordinary Revenue—						
Income tax	1,942,902	1,892,965	49,937	—	1,730,991	1,736,233
Surtax	138,600	134,700	3,900	—	132,400	131,200
Death duties	175,700	187,900	—	12,200	164,500	151,850
Stamps	70,600	75,100	—	4,500	56,600	50,250
Profits tax and excess profits tax	192,500	172,800	19,700	—	188,400	376,100
Excess profits levy	18,450	76,500	—	58,050	65,850	2,950
Other Inland Revenue duties	770	1,280	—	510	1,390	2,020
Total Inland Revenue	2,539,522	2,541,245	—	1,723	2,340,131	2,450,603
Customs	1,148,598	1,100,049	48,549	—	1,042,396	1,024,498
Excise	864,511	771,637	92,874	—	721,975	739,093
Total Customs and Excise	2,013,109	1,871,686	141,423	—	1,764,371	1,763,591
Motor vehicle duties	86,987	78,980	8,007	—	73,115	67,658
Post Office (net receipt)	—	5,100	—	5,100	—	—
Broadcast receiving licences	25,800	22,300	3,500	—	16,400	14,800
Receipts from sundry loans	29,858	23,023	6,835	—	38,014	24,681
Miscellaneous receipts	197,867	195,558	2,309	—	136,099	117,382
Total Ordinary Revenue	4,893,143	4,737,892	155,251	—	4,368,130	4,438,715
Self-balancing Revenue—						
Post Office	264,900	243,600	21,300	—	232,000	209,900
Income tax deducted from excess profits tax, post-war refunds	2,365	5,186	—	2,821	5,940	5,192

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

INDEX TO VOL. 119.

YEAR 1956

(Statistical and Current Notes are indicated by the letters S.N. and C.N. respectively.)

PAGE

ABSENCE from work . . 100, 222, 353, 489
ADAMS, Arthur Avery. Obituary notice . 493
ADDITIONS to the Library. *See* LIBRARY.
ANALYSES of covariance and variance, a geometrical derivation of. *See* Cox, C. P.
ANALYSIS of business concentration, The. *See* HART, P. E. and PRAIS, S. J.
ANNUAL General Meeting, Proceedings of 122nd, 1956 477
ANSTEY, Vera and TARAPOREVALA, Russi J. Some aspects of the structure of Indian industry 62–82

BANK of England: Weekly Return . 122–123
BANKING statistics. *See* DACEY, W. Manning.
BARTLETT, M. S. Note on a visit to Moscow for the Third Soviet Mathematical Congress 456–458
BIRTHS, deaths and marriages. *See* REGISTRATION.
BOOKS. Reviews and notices of statistical and economic books.
Abel-Smith, B. and Titmuss, R. M. The cost of the national health service in England and Wales. 214
Adams, J. K. Basic statistical concepts . 211
Allais, M. Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école Américaine 213
Annual abstract of statistics 1955. (S.N.) . 100
Barker, R. E. Books for all . . . 486
Blacker, C. P. and Gore, A. T. The Bethlem Hospital and the Maudsley Hospital triennial statistical report 1949–1951 . . 215
Braithwaite, R. B. Theory of games as a tool for the moral philosopher . . 86
Bush, R. R. and Mosteller, F. Stochastic models for learning . . . 87
Cartter, A. M. The redistribution of income in post-war Britain . . . 95
Cohen, J. and Hansel, M. Risk and gambling 211
Cornell, F. G. The essentials of educational statistics 480
Davidson, A. R. History of the Faculty of Actuaries in Scotland, 1856–1956 . . 351
Devons, E. An introduction to British economic statistics . . . 478
Dornbusch, S. M. and Schmid, C. F. A primer of social statistics . . . 341
Dudding, B. P. and Jennett, W. J. Control chart technique when manufacturing to a specification 97
Economic Commission for Europe. The European housing situation . . 349
—The problem of cost in the inland transport industry 490
Economic Survey of Europe in 1955 (S.N.) . 224
Federer, W. T. Experimental design, theory and application 340
Fieller, E. C., Lewis, T. and Pearson, E. S. Correlated random normal deviates . 345

PAGE

Finney, D. J. Experimental design and its statistical basis 210
Galenson, W. Labor productivity in Soviet and American industry . . . 96
Grünhert, Max. Juvenile offenders before the Courts 480
Guide to New Zealand official statistics (S.N.) 101
Harvard University computation laboratory. Tables of the cumulate binomial probability distribution 344
Haycocks, H. W. and Perks, W. Mortality and other investigations, Vol. I . . 88
Hickman, C. A. and Kuhn, M. H. Individuals, groups, and economic behavior . . 481
Hoffman, W. British industry, 1700–1950 . 347
Income and Wealth, Series IV. Edited by Milton Gilbert and Richard Stone . 217
Interdepartmental Committee on social and economic research, report (S.N.) . . 102
Kendall, M. G. Rank correlation methods 2nd ed. 90
Kopal, Z. Numerical analysis . . . 212
Kulldorff, G. Migration probabilities . 216
London Transport Executive. Health in industry 349
McCloskey, J. F. and Trevethen, F. N. (*ed.*). Operations research for management . 94
Mannheim, H. and Wilkins, L. T. Prediction methods in relation to Borstal training . 91
Malzburg, B. and Lee, E. M. Migration and mental disease 348
Mehta, M. M. Structure of Indian industries 218
Miller, Herman P. Income of the American people 219
Ministry of Fuel and Power statistical digest, 1955 (S.N.) 354
Ministry of Labour and National Service. Report on proposals for a new index of retail prices 346
Morton, J. E. Urban mortgage lending . 482
National Bureau of Economic Research. Capital formation and economic growth . 484
National Bureau of Standards. Table of hyperbolic sines and cosines $x = 2$ to $x = 10$ 344
—Table of the descending exponential $x = 2{\cdot}5$ to $x = 10$ 344
National Food Survey Committee, Annual report (1953) (S.N.) 100
Neville, E. H. Rectangular polar conversion tables 346
Nielsen, K. L. Methods in numerical analysis 343
Organisation for European Economic Co-operation—Demographic trends in Western Europe, 1951–1971 (S.N.) . . . 354
Ostle, B. Statistics in research . . 87
Patnaik, P. B. Training and research in statistics in India (S.N.) . . . 101
Petersen, William. Planned migration: the social determinants of the Dutch-Canadian movement 216
Political and Economic Planning. World population and resources 93
Pompilj, G. and Napolitani, D. Piano degli esperimenti 90
Registrar-Generals' Statistical review of England and Wales. Text volume (C.N.) 100
Roy, René, ed. Cahiers du seminaire d'économétrie, No. 3 486

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PAGE

Sainsbury, P. Suicide in London . . 92
Sprowle, R. C. Elementary statistics . 339
Tippett, L. H. C. Statistics. 2nd ed. . 342
United Nations Statistical yearbook (C.N.) . 223
—World economic survey for 1955 (C.N.). 490
—Yearbook of international trade statistics 223
U.S. Department of Commerce. National income (C.N.) 224
Vajda, S. The theory of games and linear programming 340
Wallis, W. A. and Roberts, H. V. Statistics: a new approach 479
Wishart, J. and Sanders, H. G. Principles and practice of field experimentation, 2nd ed. 89
World Health Organisation. Annual epidemiological and vital statistics, 1953 . 484
Woytinski, W. S. and E. S. World commerce and government 350

BUSINESS concentration. *See* HART, P. E. and PRAIS, S. J.

CAUSAL inference from observational data. *See* WOLD, Herman.

COLE, Dorothy and UTTING, J. E. G. Estimating expenditure, saving and income from household budgets . 371–387

The survey methods 371
Principal results 374
Sources of error 377
Non-statistical biases . . . 379
Estimates—expenditure, income, saving . 380
Conclusions 385
Appendix: main contents of the commodity groups 386

Discussion: Professor Allen; Mr. Saunders; Mr. W. F. F. Kemsley; Miss M. Deane; Mr. Durbin; Mr. Mason; the authors, in reply . . . 387–392

COUNCIL, Report of, for the financial year ended December 31st, 1955 and for the sessional year ended June 20th, 1956 459–476

COX, C. P. A geometrical derivation of the analyses of covariance and variance 333–335

DACEY, W. Manning. Banking statistics (sources and nature of official information) 199–209

ELECTIONS to Fellowship 27, 61, 149, 191, 283, 311, 392, 470, 477

EMMETT, B. P. The television audience in the United Kingdom . . . 284–306

Size of television public . . . 284
Nature of the television public . . 293
Viewing of the television public . . 298
Guest viewing 305
Summary 306

Discussion: Mr. A. Stuart; Mr. H. F. Lydall; Professor P. E. Vernon; Mr. E. J. Davis; Mr. A. Gabor; Mr. Wood; Dr. Barna; Mr. J. A. Bound; Mr. J. I. Mason; Mr. Emmett in reply 306–311

EMPLOYMENT 99, 222, 352, 488
———. Age analysis of employees . 353

ESTIMATING expenditure, saving, and income from household budgets. *See* COLE, Dorothy and UTTING, J. E. G.

FACTOR analysis, principles of. *See* KENDALL, M. G. and LAWLEY, D. N.

PAGE

FIELD experiments. The statistician and the planning of. *See* FINNEY, D. J.
FINANCES of the Society . . . 468
FINNEY, D. J. The statistician and the planning of field experiments . . 1–17

Nature of experimental agriculture . . 2
Statistics and scientific research . . 3
Design of experiments . . . 5
Statistical analysis—Reporting and interpreting 14

Discussion: Dr. Yates; Dr. S. C. Pearce; Professor A. Bradford Hill; Mr. G. H. Freeman; Mr. G. M. Clarke; Dr. Ruben; Mr. F. J. Anscombe; Dr. Vajda; Mr. J. A. Nelder; Dr. Sampford; Mr. Skellam; the Chairman; Mr. C. P. Cox; Dr. Finney, in reply . 17–27

FISHER, Sir Ronald, F.R.S. Award of Copley Medal 461

GEOMETRICAL derivation of the analyses of covariance and variance. *See* COX, C. P.
GEOMETRY of statistics, Some aspects of the. *See* PEARSON, Professor E. S.
GLENDAY, R. Richard Jones: a reappraisal 192–193

HART, P. E. and PRAIS, S. J. The analysis of business concentration: a statistical approach 150–181

Methods of measuring business concentration 152
Comparative statics 154
Dynamics of business concentration . . 161
Conclusions 175
Appendices 176

Discussion: Professor Champernowne; Mr. Gorman; Mr. Colin Clark; Professor M. G. Kendall; Mr. Mason; Mr. Bluett; Dr. T. Barna; the authors in reply 181–190

HISTORY of sampling methods in Russia. *See* ŽARKOVIĆ, S. S.
HONOURS. Awards to Fellows in 1956 . 101
HOUSEHOLD budgets, estimating expenditure, savings, and income from. *See* COLE, Dorothy, and UTTING, J. E. G.
HOUBLON-NORMAN Fund, offer of Fellowships, 1957; awards, 1956 . . 491

INDEX numbers, some theory of. *See* SENG, You Poh.
INDIAN industry, structure of. *See* ANSTEY, Vera and TARAPOREVALA, Russi J.
INDICES of wholesale prices. *See* PHILLIPS, H. S.

KENDALL, M. G. and LAWLEY, D. N. The principles of factor analysis . . 83–84

LIBRARY, Additions to . 112, 231, 361, 500
——, Report on 467

MEDICAL Section. Organisation, Membership, the Section Committee . . 85
METHORST, H. W. Obituary notice . 102

NEW ZEALAND agricultural sample survey. *See* TAYLOR, W. B., and CLEMENT, D. V. P.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PAGE

OBITUARY:—
Arthur Avery Adams 493
H. W. Methorst 102
Charles Stewart Orwin 103
Mrs. J. R. N. Stone 494
John Wishart 492
See also (Council report) . . 459

OPERATIONAL research, international conference on 490
ORWIN, Charles Stewart. Obituary notice 103

PEARSON, Professor E. S. Some aspects of the geometry of statistics. Inaugural address of the President, delivered Jan. 18th, 1956 125–146
Geometry and teaching; graphical approach; geometry of two and three dimensions; diagrams showing component parts in analysis of variance 127
Seeing the data as a whole; diagrams and textbooks 137
Concluding remarks 143
Proceedings of the Meeting: Lord Piercy; Professor Bradford Hill; the President, in reply 146–149

PHILLIPS, H. S. United Kingdom indices of wholesale prices, 1949–1955 . 239–276
Part I. A description of the new scheme of Board of Trade wholesale price index numbers (June 30th, 1949 = 100) and of the practical problems surmounted in constructing and maintaining them . . 240–256
Part II. The relative movements of the main wholesale price index numbers in the new series: some analyses and comparisons . 256–270
Appendices: Problems met in obtaining co-operation of industry. Storing and processing price information; calculating price relatives and index numbers . . . 271–276
Discussion: Professor C. F. Carter; Mr. H. S. Booker; Mr. Stafford; Miss M. Deane; Professor Allen; Dr. Barna; Mr. G. R. White; Mr. Leak; Dr. J. Wise; Dr. Phillips in reply . . 276–283

PRICES, Retail . . . 99, 221, 352, 488
—— New Ministry of Labour Index . 221
PRICES, Wholesale, in 1955. *See STATIST, editor of.*
PRICES, Wholesale, indices of. *See* PHILLIPS, H. S.
PRINCIPLES of factor analysis. *See* KENDALL, M. G. and LAWLEY, D. N.

REGISTRATION of the United Kingdom and Eire 367–370
RESIDENCE and Workplace. Report on, Census, 1951 353
REYNOLDS, D. J. The cost of road accidents 393–408
REVENUE of the United Kingdom, 1955 . 505
RICHARD JONES: a reappraisal. *See* GLENDAY, R.

PAGE

SAMPLING Methods in Russia, on the history of. *See* ŽARKOVIĆ, S. S.
SECTIONS and Groups of the Society . 462–466
SENG, You Poh. Some theory of index numbers. Part I 312–332
—— Part II 425–455
SOCIAL and economic research, Standing interdepartmental committee on . 489
SOME aspects of the geometry of statistics. *See* PEARSON, Professor E. S.
SOME aspects of the structure of Indian industry. *See* ANSTEY, Vera and TARAPOREVALA, Russi J.
SOME theory of index-numbers. *See* SENG, You Poh.
Statist, The, Editor of. Wholesale Prices in 1955 194–198
Method of construction of the index numbers 194
Annual and monthly index-numbers, 1902–11 to 1955 194–195
Average price of each of the 45 commodities . 196–197
Index of silver prices—gold and silver production 198

STATISTICIAN, The, and the planning of field experiments. *See* FINNEY, D. J.
STATISTICAL and economic articles in recent periodicals . . . 105, 225, 355, 495
STONE, Mrs. J. R. N. Obituary notice . 494

TAYLOR, W. B. and CLEMENT, D. V. P. The New Zealand agricultural sample survey 409–424
TELEVISION audience in the United Kingdom. *See* EMMETT, B. P.
TRADE. World trade in 1954 (U.N. figures) 223
—— Union membership . . . 100

UNEMPLOYMENT figures . 100, 222, 352, 488

WAGE rates and average earnings 99, 221, 352, 488
WISHART, John. Obituary notice . . 492
WOLD, Herman. Causal inference from observational data . . . 28–50
1. A fourfold map of ends and means in applied statistics 29
2. A review of explanatory approaches, with particular regard to the treatment of experimental *v.* observational data. . 34
3. The explanatory analysis of observational data: some typical situations . . 41
Discussion: Professor Richard Stone; Mr. J. Durbin; Professor P. A. P. Moran; Dr. Klein; Professor Bartlett; Mr. F. E. A. Briggs; Professor M. G. Kendall; Professor Wold in reply 51–60

ŽARKOVIĆ, S. S. Note on the history of sampling methods in Russia . . 336–338

पुस्तकालय
110637
गुरुकुल कांगड़ी विश्वविद्यालय
सन्दर्भ ग्रन्थ
REFERENCE BOOK

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

AGE

-466

-332
-455

489

198

195
197

495
494

424

223
100

488

488
492

-60

-60

38

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

सन्दर्भ ग्रन्थ
REFERENCE BOOK

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

सन्दर्भ ग्रन्थ
REFERENCE BOOK

Compiled
1999-2000

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar